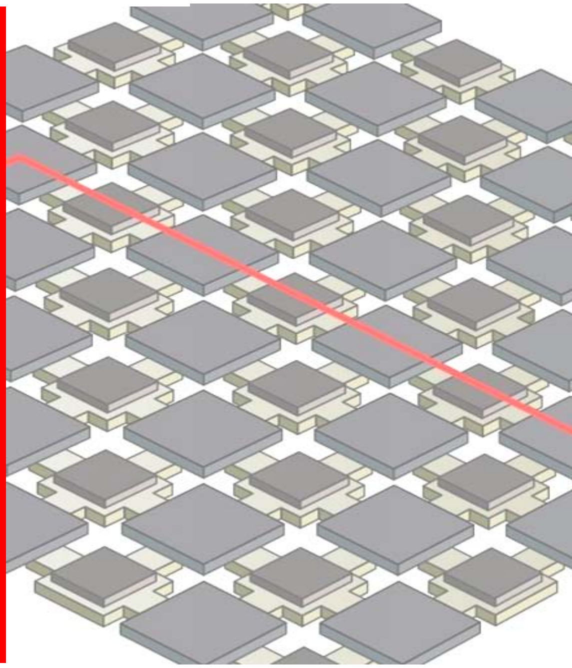
ORACLE



Acknowledgements:

J. Cunningham, R. Ho, X. Zheng, J. Lexau, H. Thacker, J. Yao, Y. Luo, G. Li, I. Shubin, F. Liu, D. Patil, K. Raj, and J. Mitchell

M. Asghari KOTURA

T. Pinguet LUXTERA

"Overview of short-reach optical interconnects: from VCSELs to silicon nanophotonics"

Ashok V. Krishnamoorthy (Jack Cunningham)
Hardware Architect
Sun Labs, Oracle
Physical Sciences Center, San Diego, CA

This work was supported in part by DARPA under contract NBCH30390002. The views, opinions, and/or findings contained in this article/presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense

ORACLE 30 ORACLE THIRTY YEARS

Outline

Introduction

- Definitions
- Penetration of optics into communication systems

Fibers, connectors, and module packaging

- Optical product segmentation
- Some examples of systems using optical interconnects

Optics to the package/chip

Link energy efficiency metric and goals

Silicon photonics and WDM

Overview of recent optical component results

Brief introduction to the macrochip

ORACLE

Optical Transceivers

> Integrated modules incorporating optical laser transmitters and photodiode receivers. These modules convert physical signals from electrical to optical and vice-versa in a network and couple the optical signals into (and out of) optical fiber. Transceivers have serial electrical interfaces on the host board.

Parallel Optical Transceivers, Modules, Interconnects or "Parallel Optics"

> Integrated optical laser transmitter and receivers incorporating multiple signaling channels in a single housing, each channel having a separate serial electrical interface to the host board. Typical values are 12 channels, although higher numbers (24, 36) have been developed. Parallel optical modules typically utilize an array of VCSELs and detectors to transmit and receive optical signals traveling in multi-mode fibers over a distance of up to 300m.

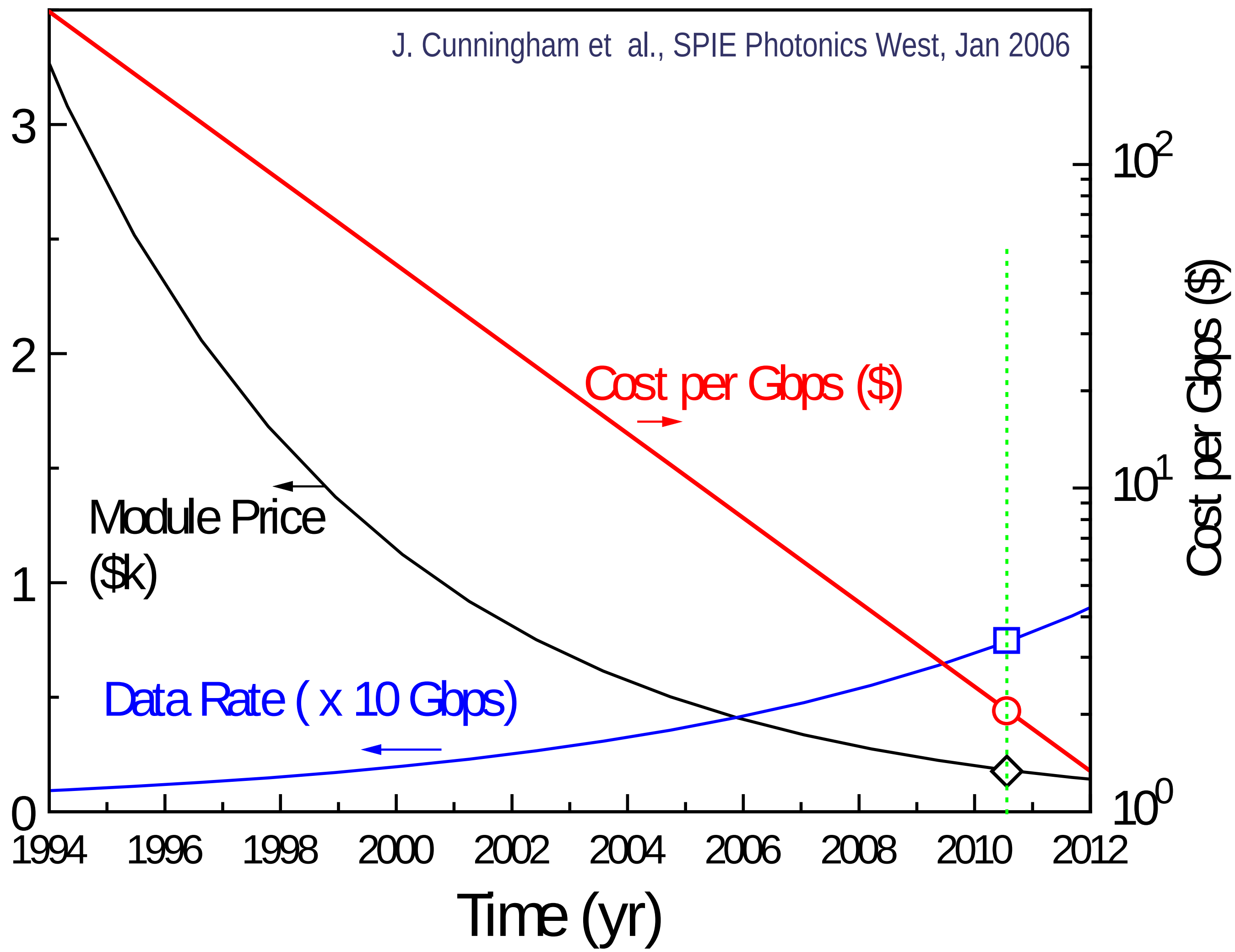
VCSEL

> Is a type of semiconductor laser diode with laser beam emission perpendicular from the top surface, contrary to conventional edge-emitting semiconductor laser which emit in-plane from surfaces formed by cleaved facets. VCSELs are today the most-efficient, lowest-cost, and most widely used laser source for interconnects.

WDM

> Wavelength Division Multiplexing. Enables multiple data streams of varying wavelengths ("colors") to be combined into a single fiber, significantly increasing the overall capacity of the fiber and of the connector. There are two types of WDM architectures: Coarse Wavelength Division Multiplexing (CWDM), typically handling up to 8 wavelengths, and Dense Wavelength Division Multiplexing (DWDM), supporting up to 160 wavelengths.

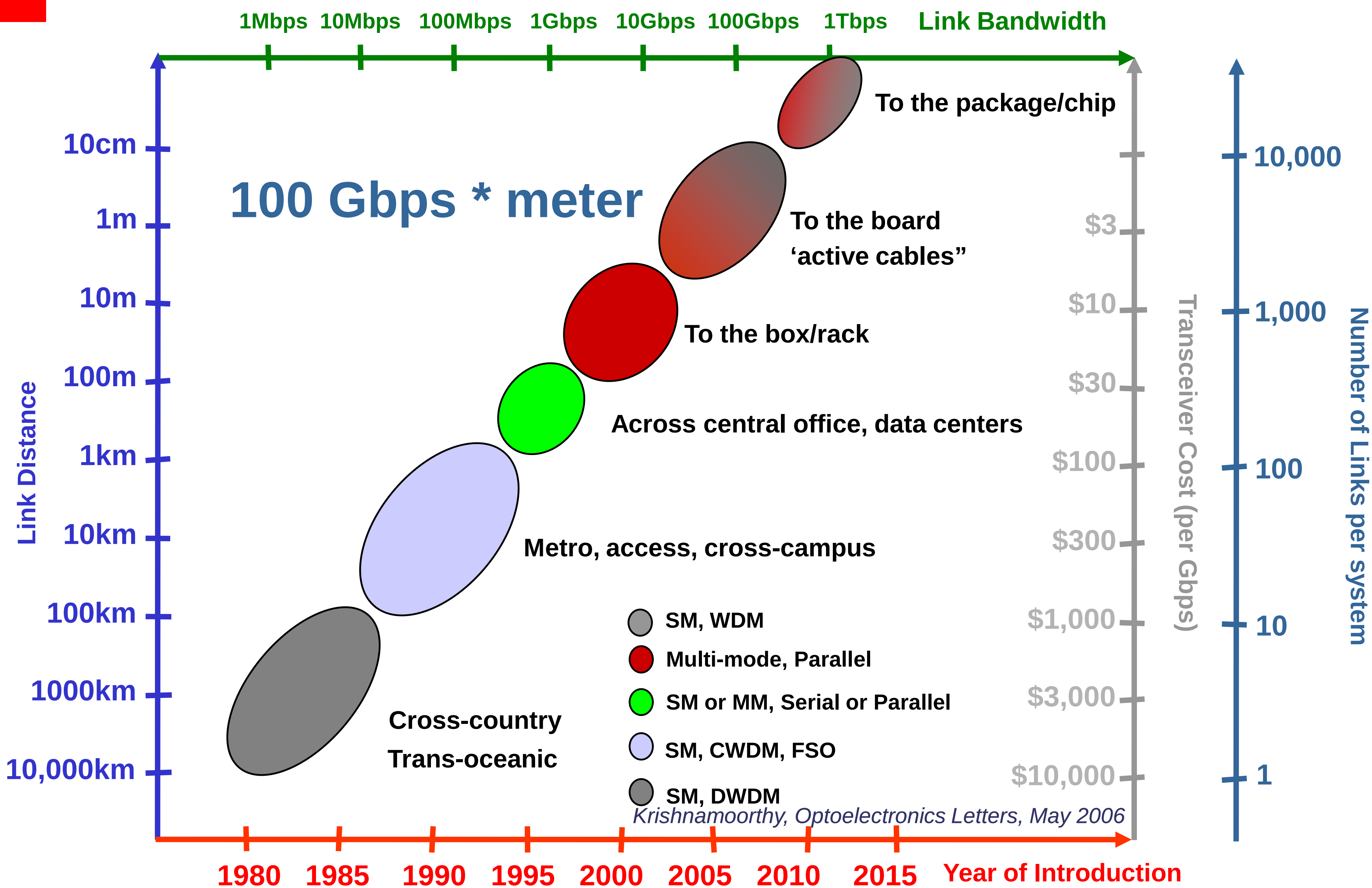
Price evolution of optical links



Approaching ~$1/Gbps

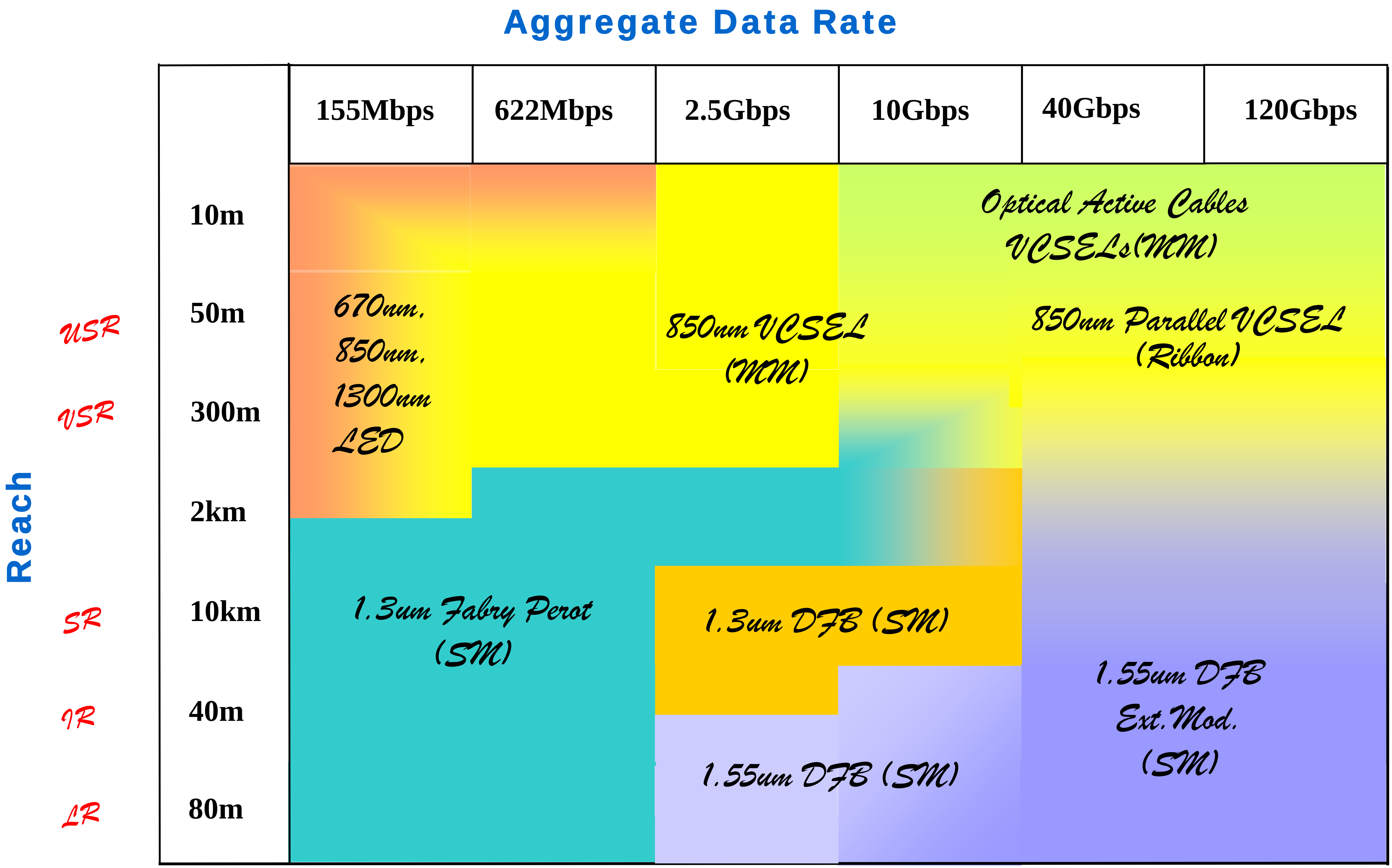
Consumer application could further reduce price

Optics in communications

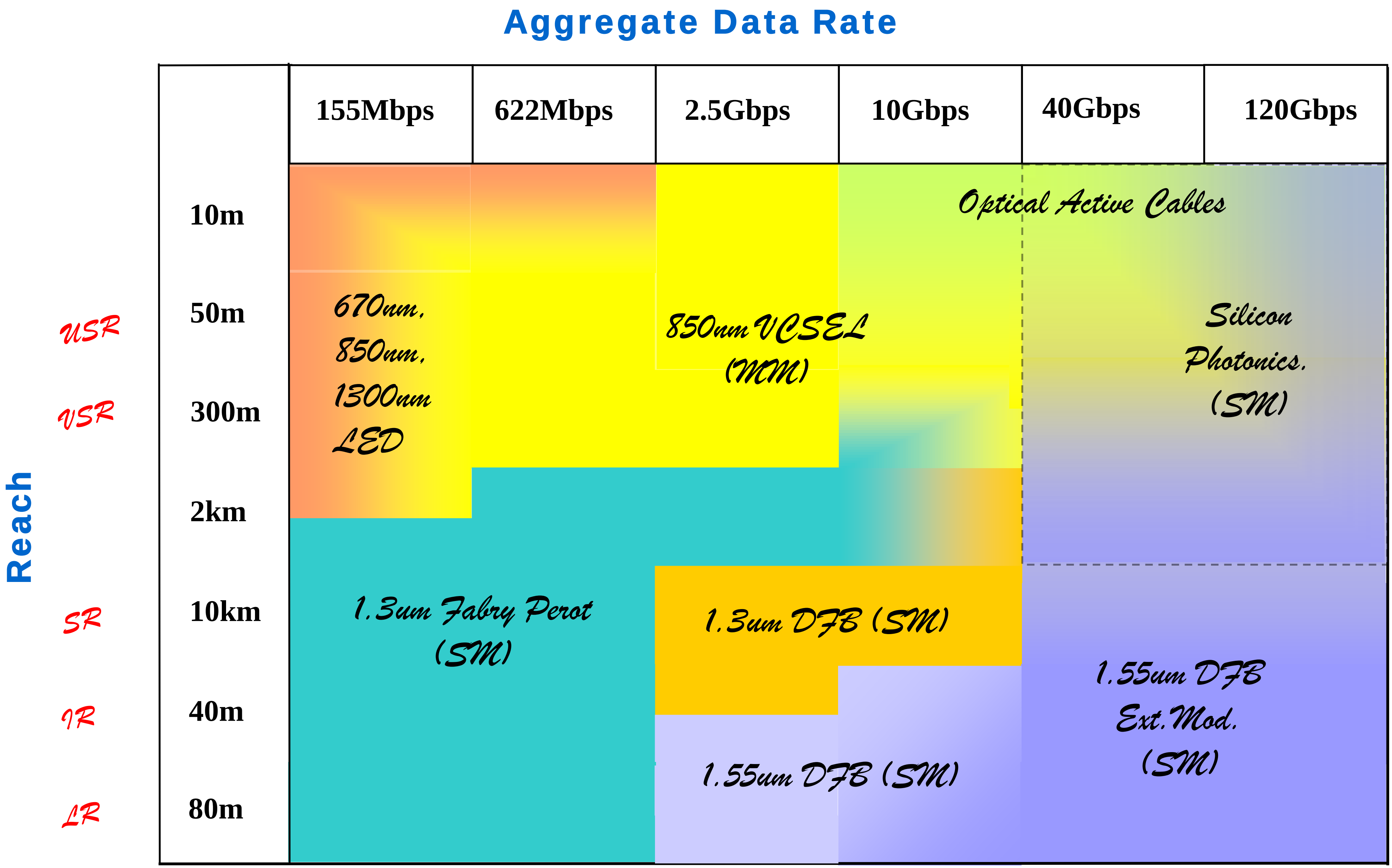


ORACLE

Optical link market segments



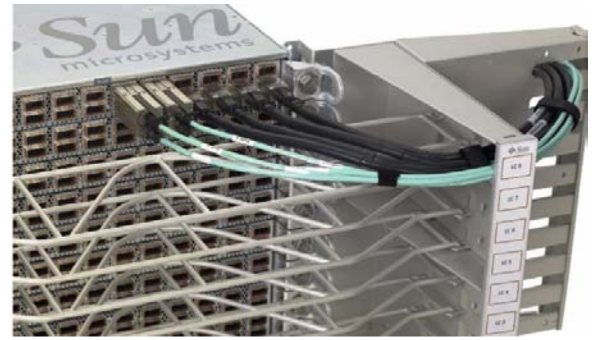
Silicon photonic interconnects



I/O for the world's largest IB switch

Gen 2: Up to 648 QDR Infiniband (40Gbps) ports [Gen 1: 3,456 SDR ports]

11 U



http://www.oracle.com/us/products/servers-storage/networking/infiniband/031556.htm

First 12x QDR cable developed by Merge Optics

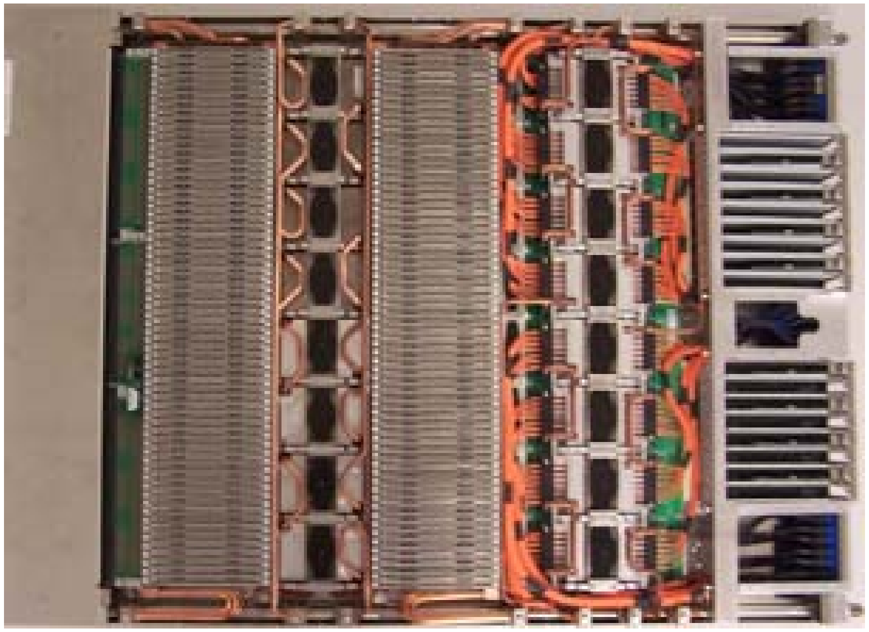
> Very high panel density requirement for Sun/Oracle QDR switch
> CXP active optical cable with three 4x10Gb/s (120Gbps per direction)
> Over 50Tbps front side I/O => Areal connection density > 1.7Tbps/sq. in

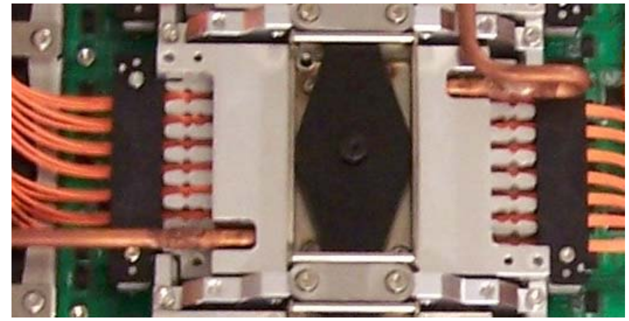
~6.8 Petabits/s of data bandwidth deployed

> Over 28,000 air-cooled VCSEL-based active cables installed
> Over 500km of these active optical cables into datacenters

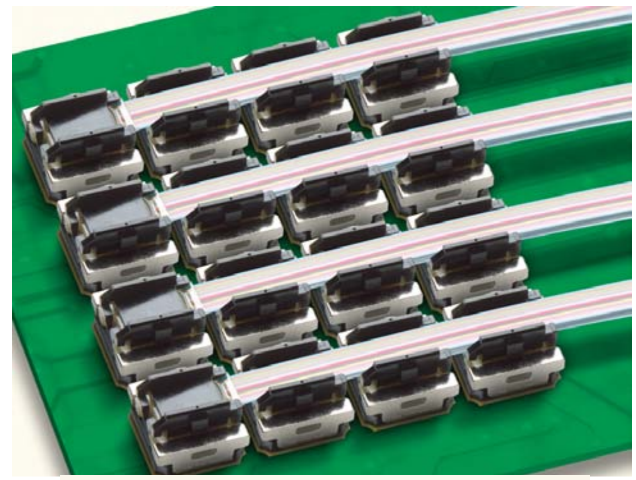
I/O for IBM’s P7-IH computing system

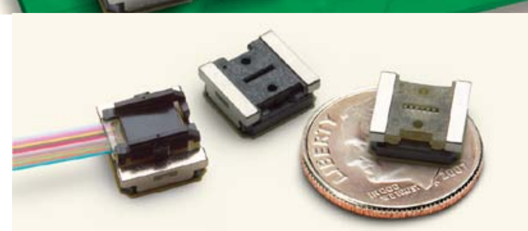
12 drawers, 8 MCMs per drawer, 4 P7 chips per MCM, 8 cores per P7





A. Benner et al., paper OTuH1, OFC 2010





Over 35Tbps of optical I/O per drawer

Each drawer can be configured as 256-way SMP

Water-cooled VCSEL modules for drawer I/O

> **areal connection density of 1.2Tbps/sq. inch**

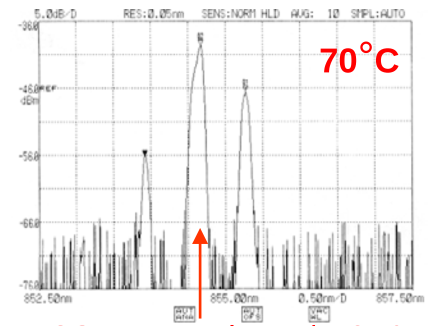
http://www.avagonow.com/Newsletters/PDFs/IEEE-MitchFields-march-021610.pdf

VCSELs and detectors on CMOS

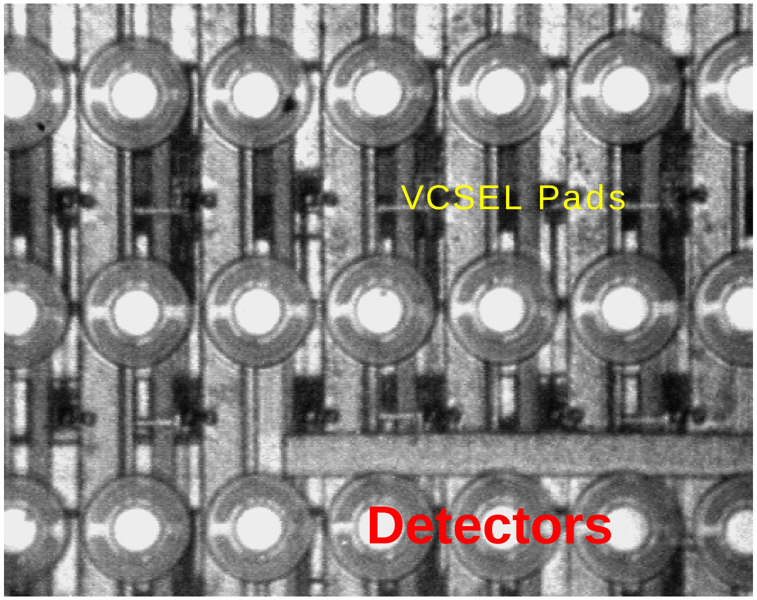
Areal density: demonstrated over 37Gbps/sq. mm (24Tbps/sq. in)

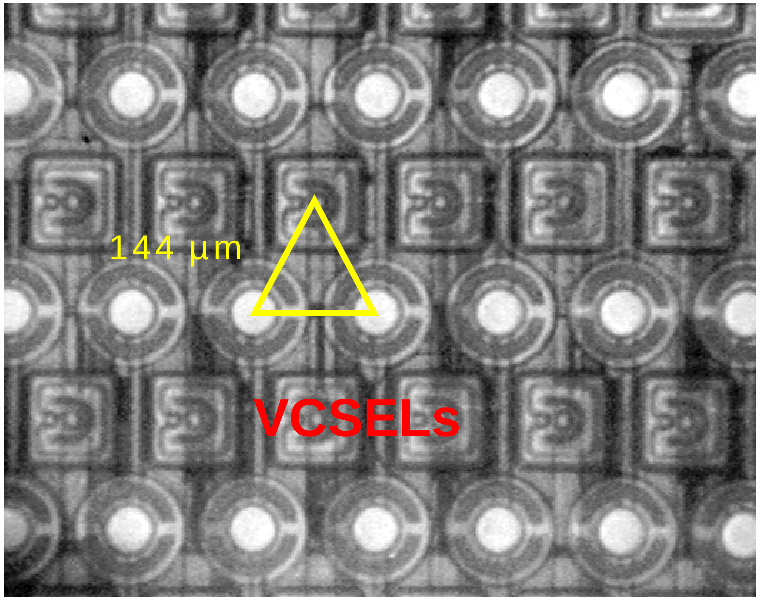
Many independent R&D efforts, e.g.

- Bell Labs - late '90s — Krishnamoorthy et al., *IEEE PTL,* August 2000
- AraLight/Xanoptix - 2002 — C. Cook et al., *IEEE JSTQE,* March/April 2003; J. Trezza et al., *IEEE Commun. Mag.,* Feb 2003
- Agilent – 2004 — B. Lemoff et al., *OSA/IEEE JLT,* September 2004
- IBM - 2009 — C. Schow et al., *OSA/IEEE JLT,* April 2009

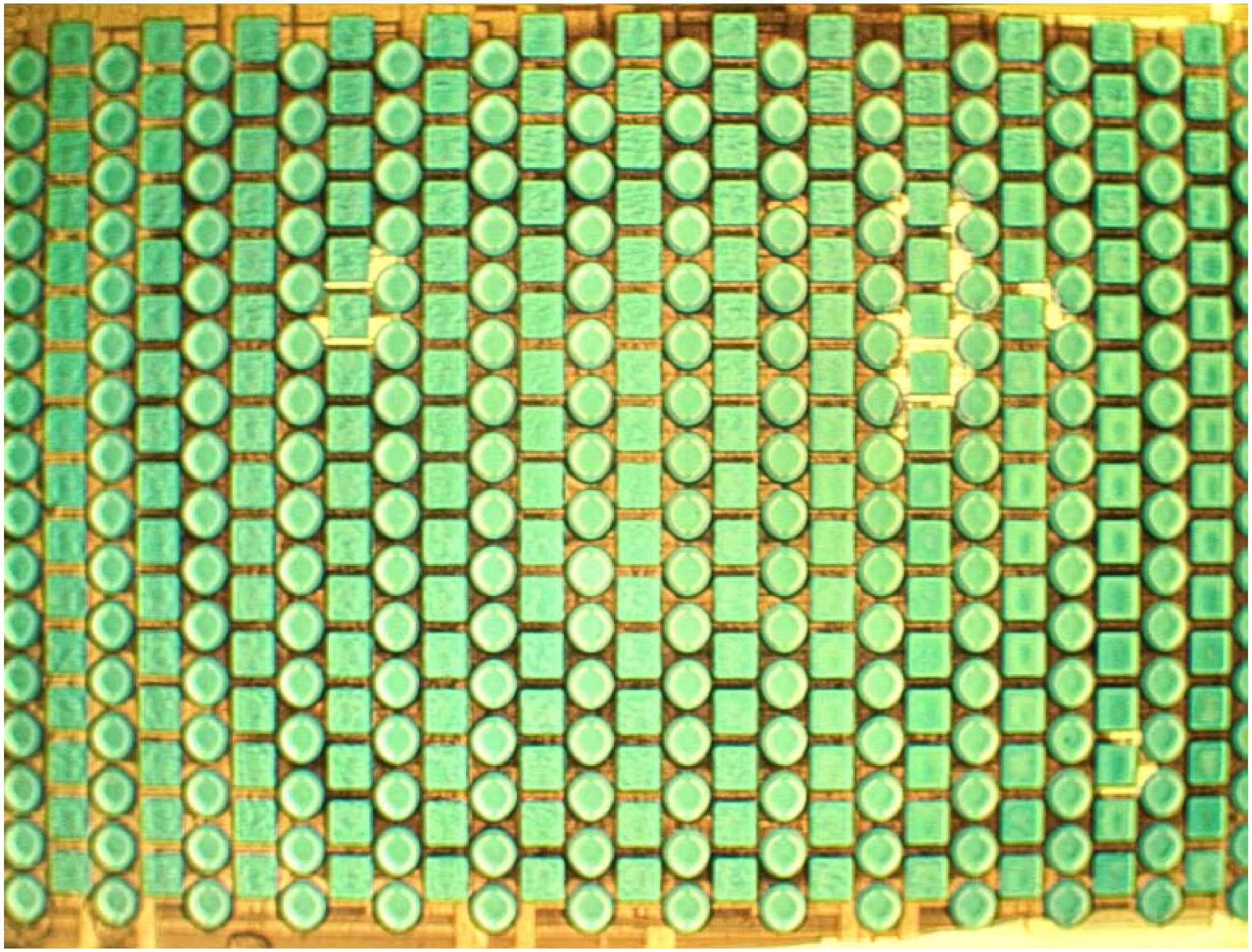


VCSEL wavelength: 850nm
(other work at 980nm)



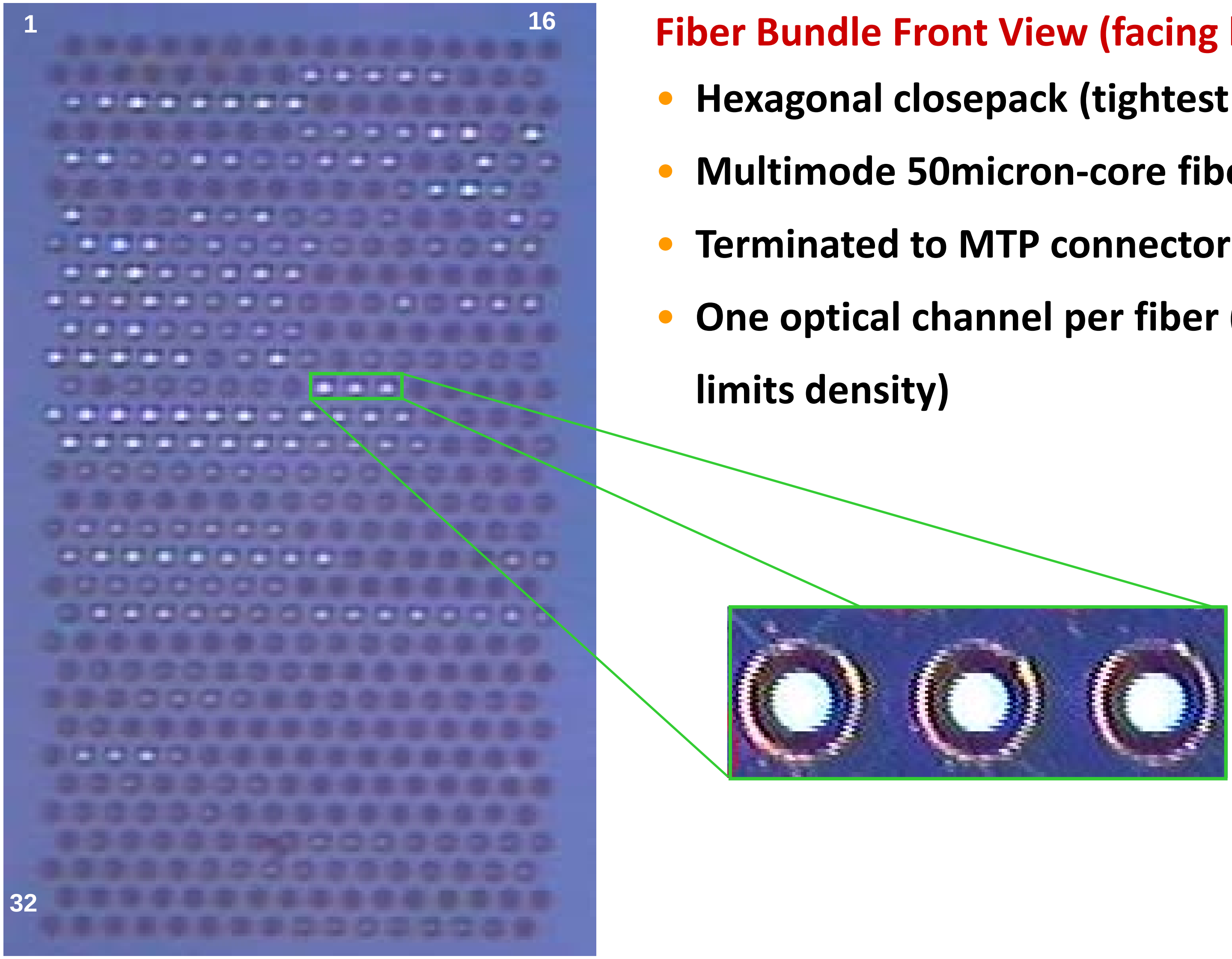


GbE switch with VCSELs & detectors



Krishnamoorthy et al., *IEEE JSTQE Spec. Issue on Green Photonics,* to appear

Multimode fiber bundle array

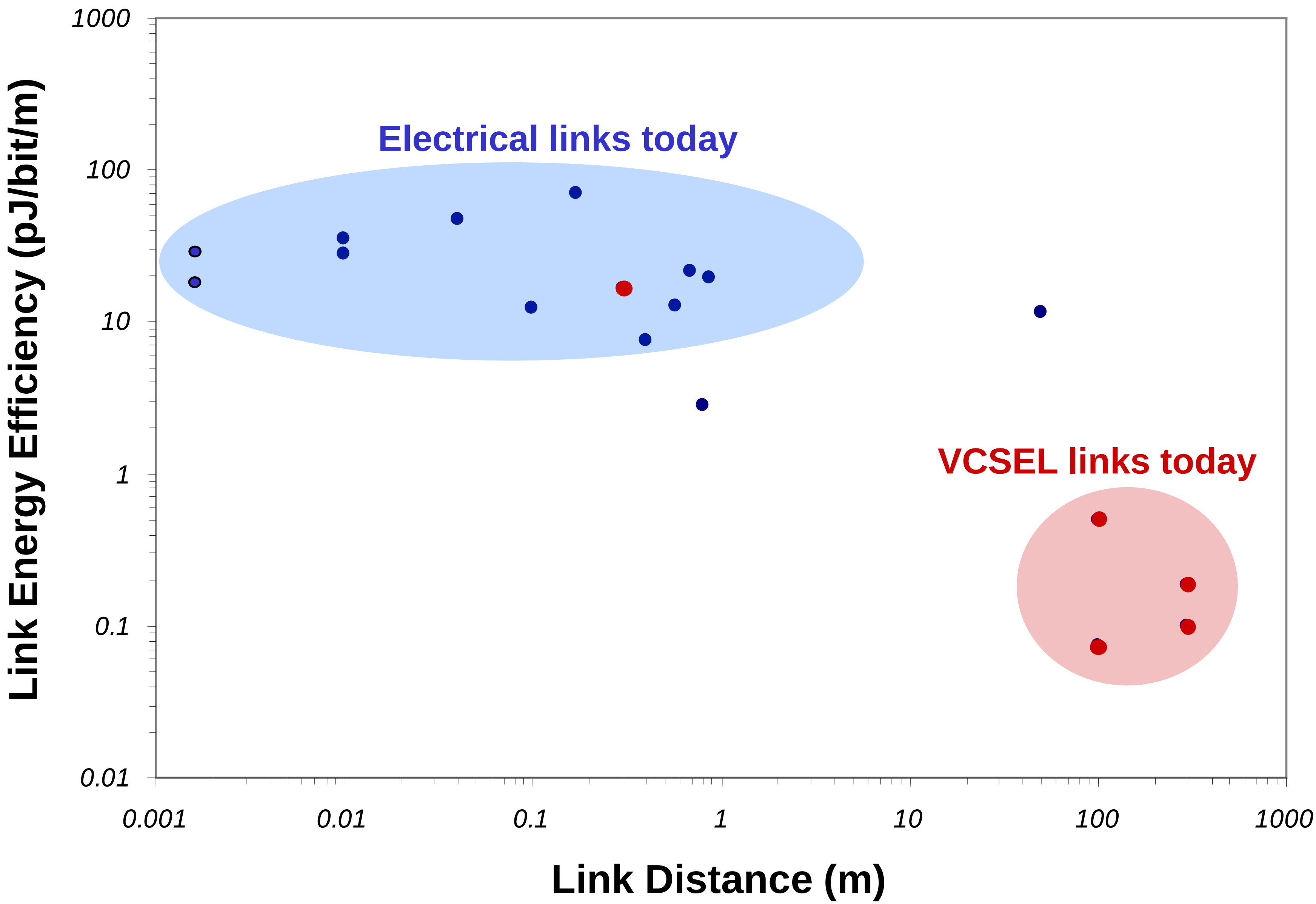


Krishnamoorthy et al., *IEEE JSTQE Spec. Issue on Green Photonics*, to appear

Fiber Bundle Front View (facing bundle)

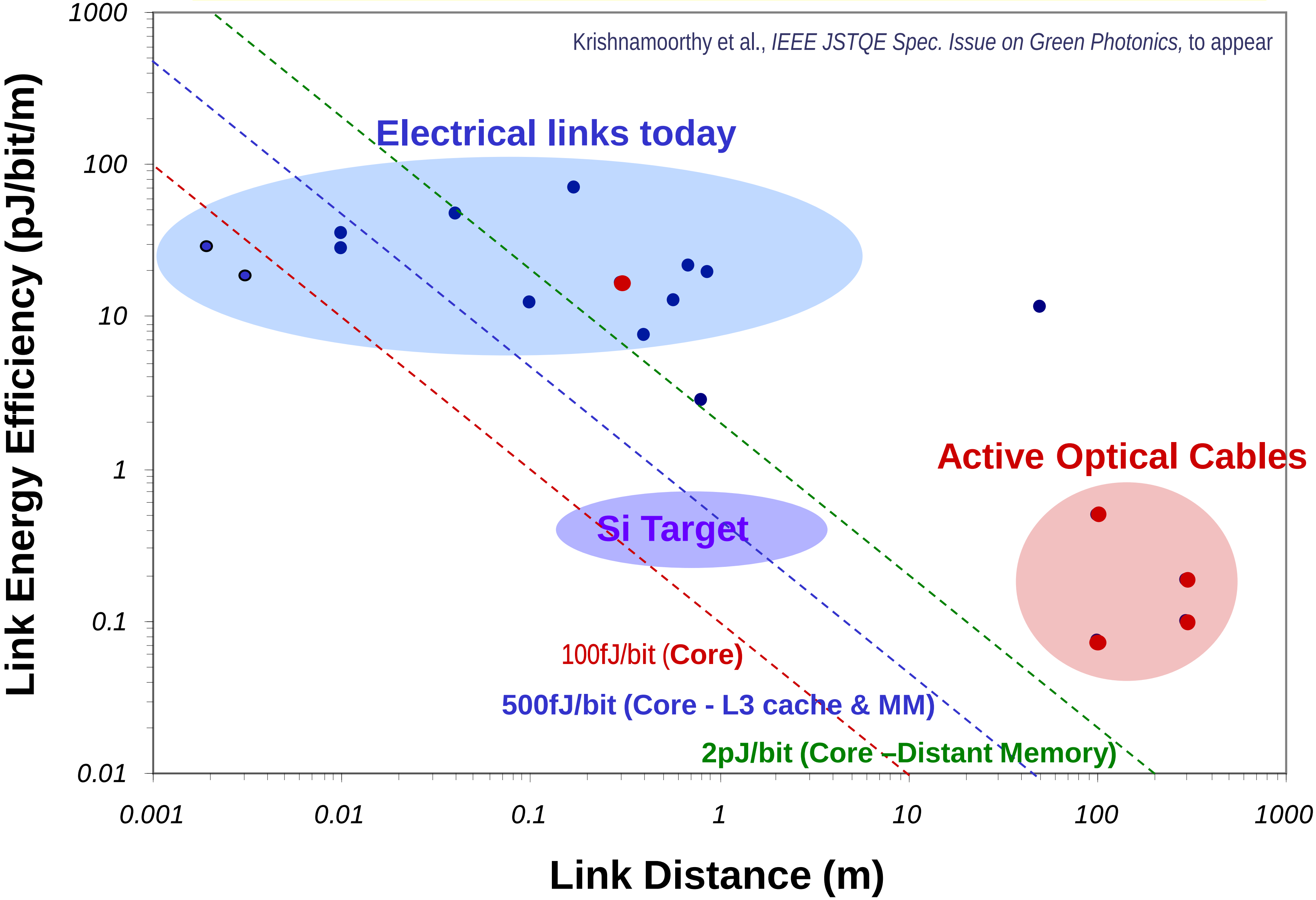
- **Hexagonal closepack (tightest geometry)**
- **Multimode 50micron-core fiber**
- **Terminated to MTP connectors at other end**
- **One optical channel per fiber (ultimately limits density)**

Link energy efficiency vs distance



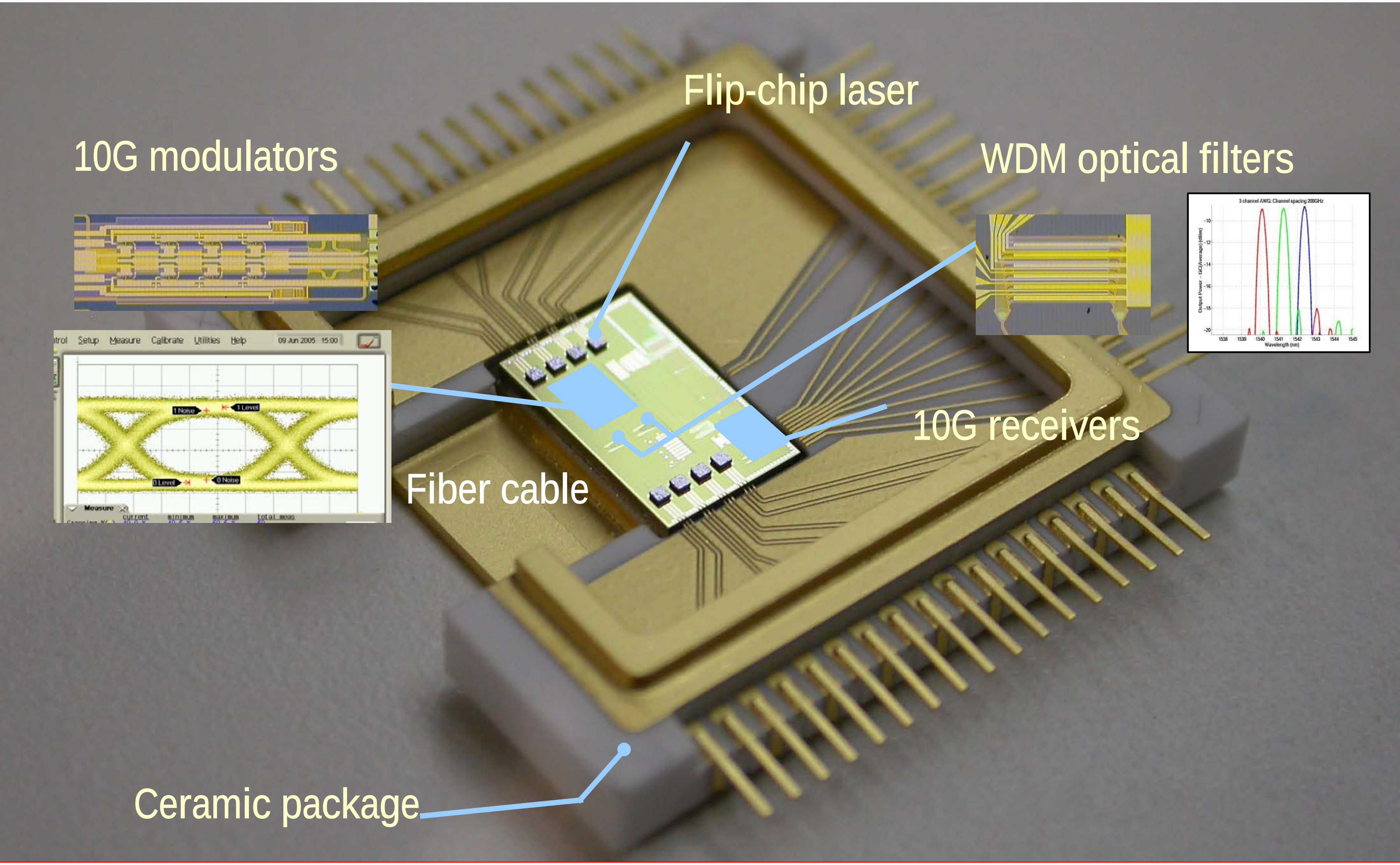
Improving link energy efficiency

System Target: <<1 mW per Gigabit/s per meter

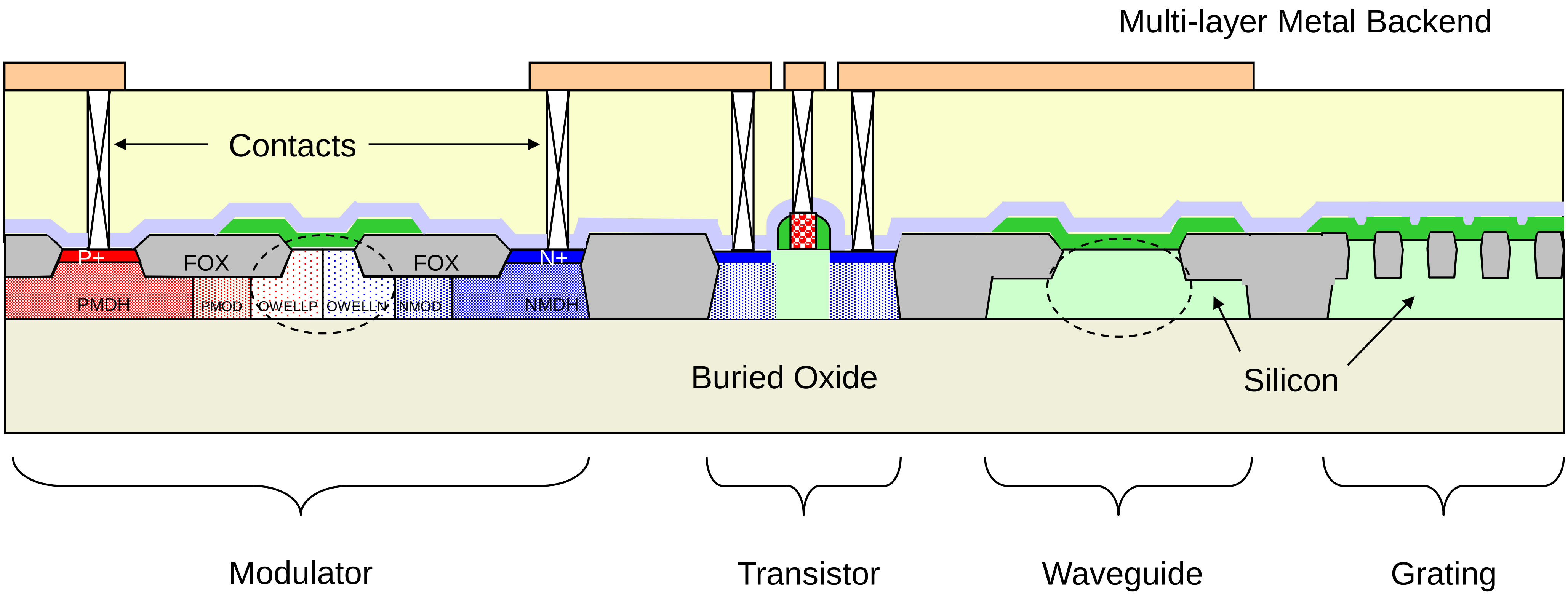


Optics to the chip: CMOS photonics

C. Gunn, IEEE Micro, March/April 2003



Introduction to CMOS photonics



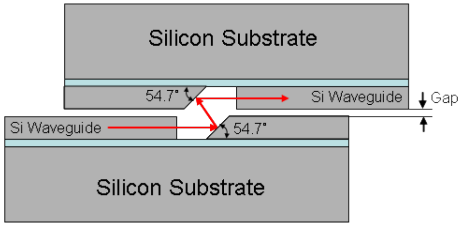
Standard silicon process with SOI wafers (e.g. Luxtera)

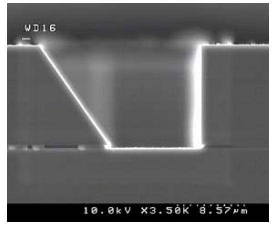
High index contrast => sub-micron structures => fast, compact devices

Proven CMOS-compatible germanium waveguide detectors

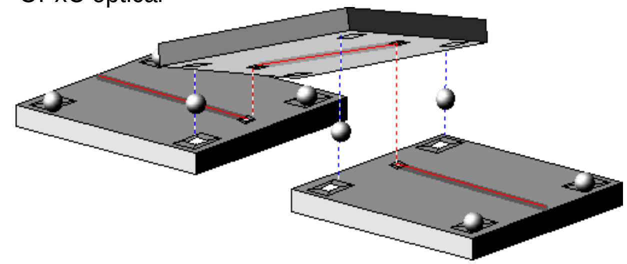
C. Gunn, IEEE Micro, March/April 2003

Optical proximity communication (OPxC)





OPxC optical



Zheng et al, *Optics Express*, September 2008,

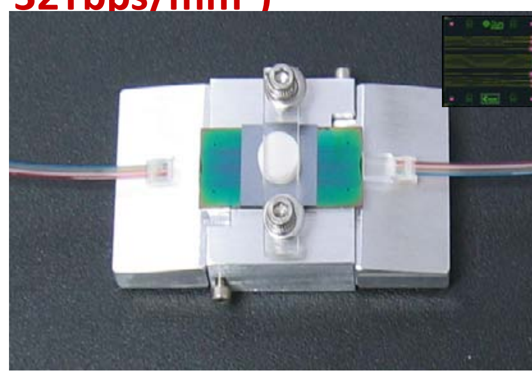
Krishnamoorthy et al., *IEEE Journal of Quantum Elec.*, July 2009

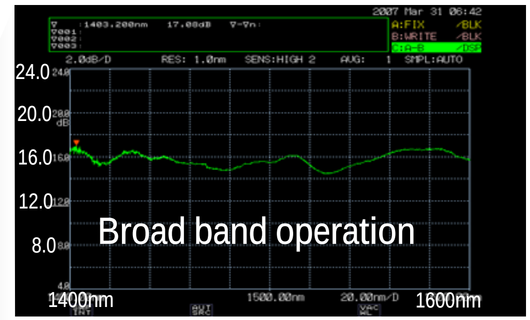
OPxC enables seamless multi-chip optical interconnects

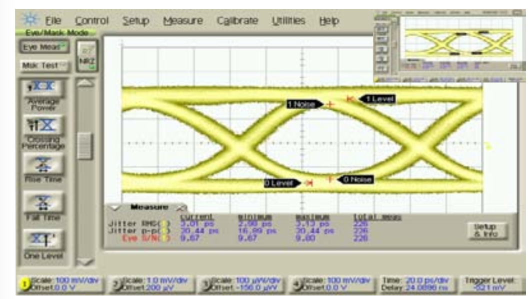
- Various approaches
  - Grating couplers, reflecting mirrors, ball lens in pit…
- High performance
  - High bandwidth density (potentially > 32Tbps/mm^2)
  - Passive coupling (no conversion pwr)
  - Performance limited by transceivers

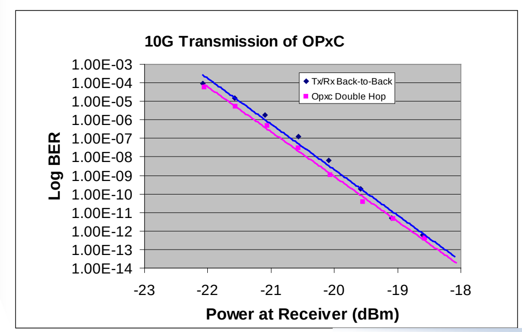
OPxC demonstration

- Reflecting mirror OPxC
  - 3 chips with 2 OPxC hops
- Promising optical performance
  - Passive alignment with etch pits and balls
  - Broad band coupling, >100nm
  - <4dB insertion loss per coupling interface
  - Negligible power penalty at receiver for 10Gbps transmission



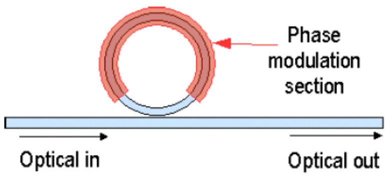


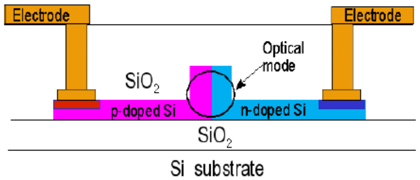


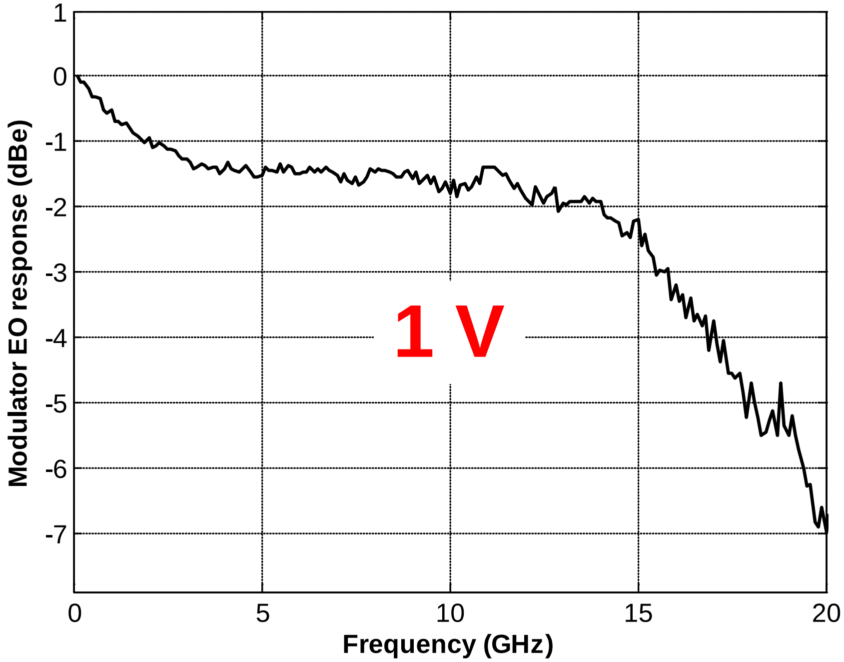


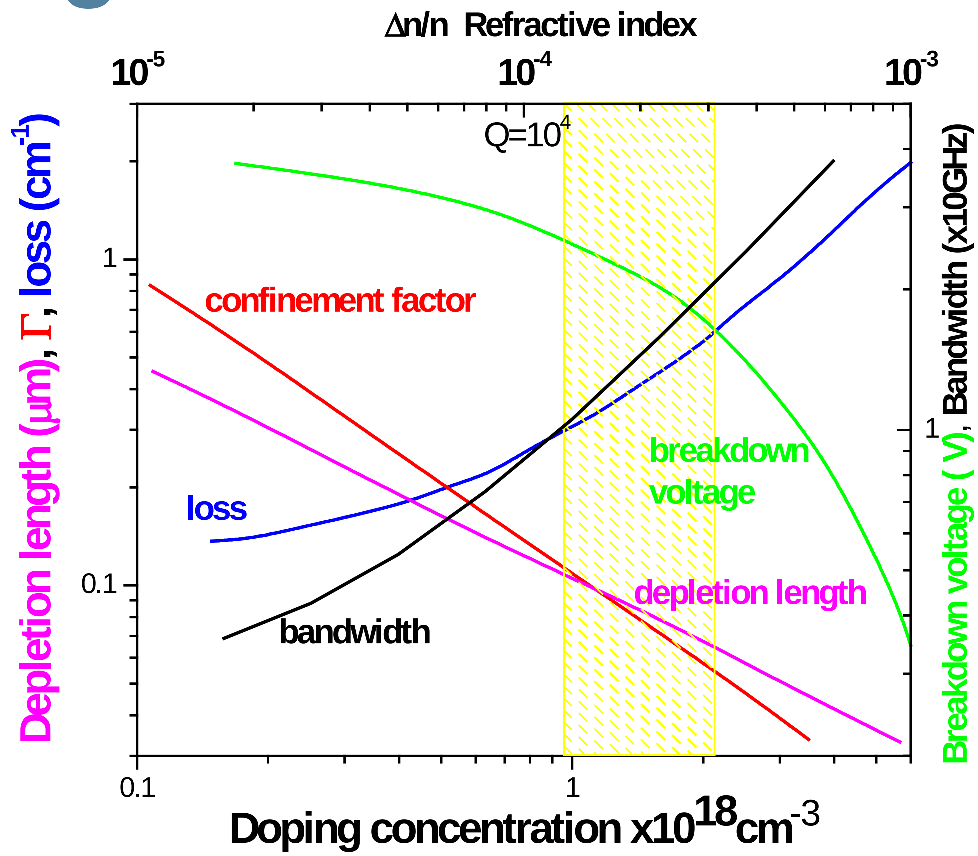
ORACLE KOTURA

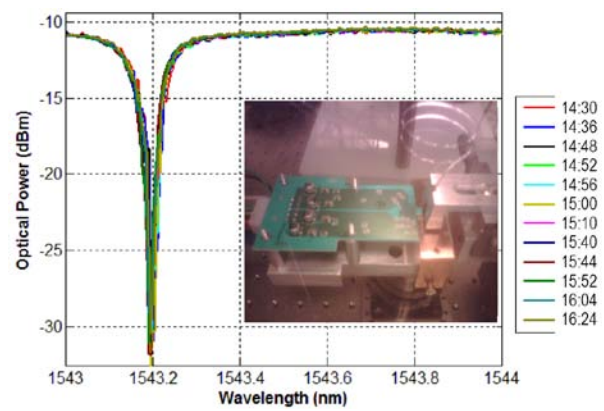
Carrier depletion ring modulator











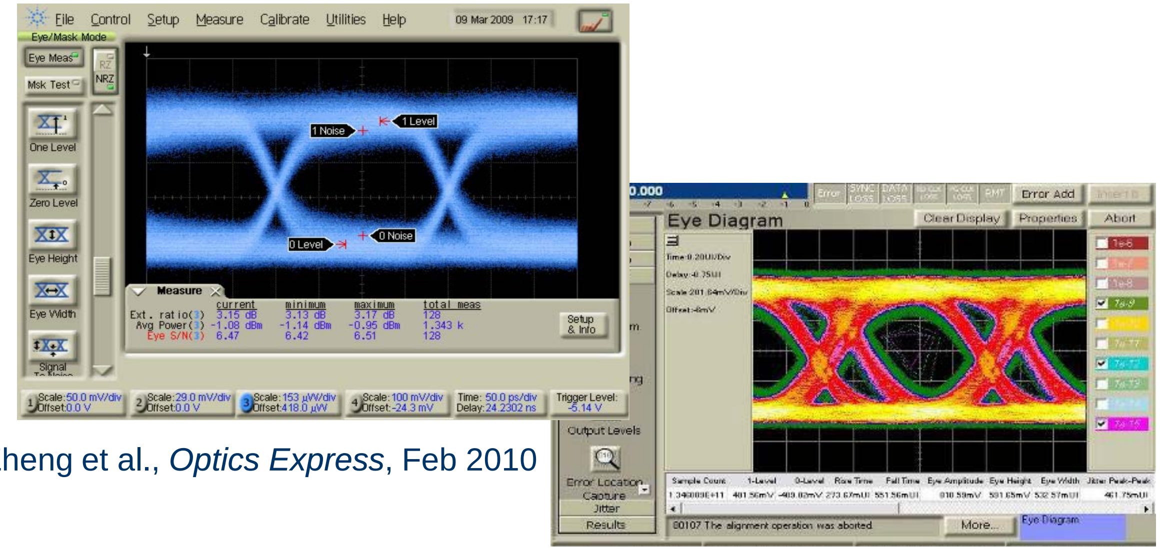
- Carrier depletion ring for low power, high speed modulation
- Free-scale 130nm SOI CMOS process
- Relatively low Q design (<10^5)
- >15GHz small-signal bandwidth with 1V reverse bias
- Stable large-signal operation (no feedback control or dynamic tuning)

X. Zheng et al., *Optics Express*, Feb 2010

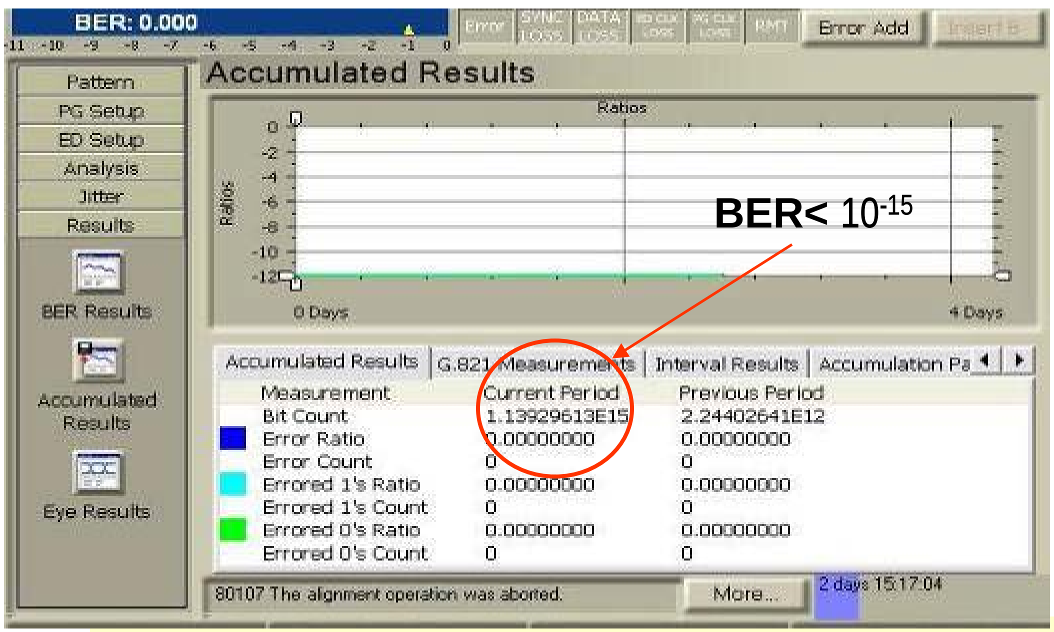
400fJ/bit all-CMOS Tx (circuits + device)

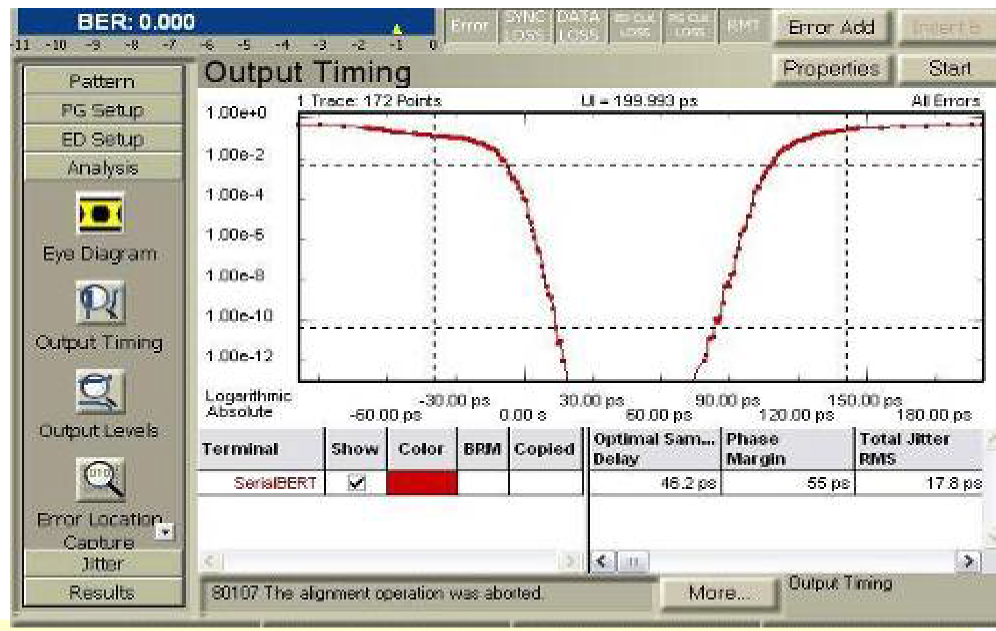
Performance Summary:

- 5Gbps, digitally clocked
- 2V, 1.95mW or 395fJ/bit
- ER 3dB; IL 6dB
- Error free transmission for over 1.5 peta bits of data
- Better than 10^{-15} BER



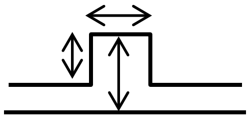
X. Zheng et al., *Optics Express*, Feb 2010

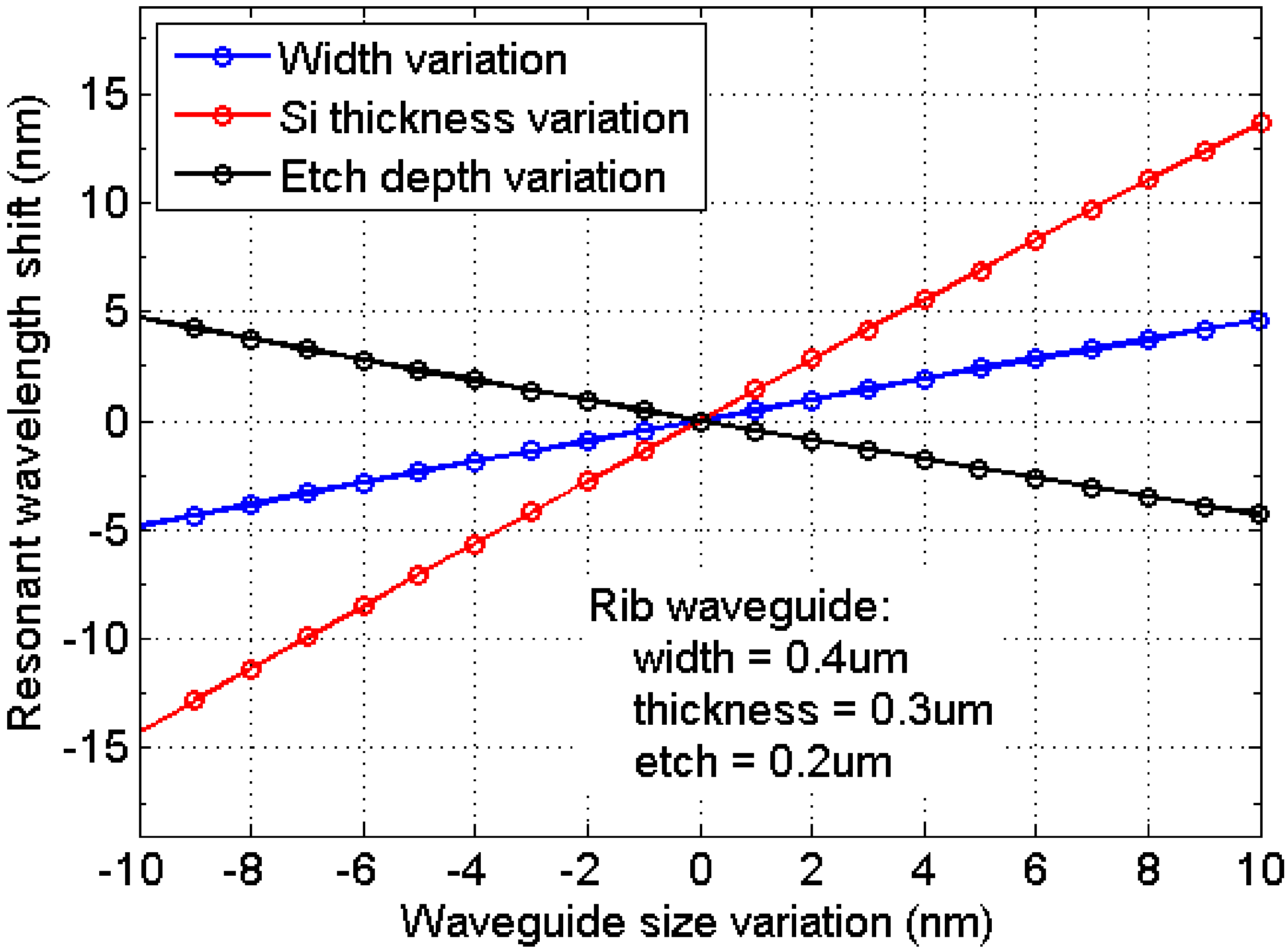




BER better than 10^{-15} with clocked digital Tx using ring modulator

Tuning out resonance imperfections

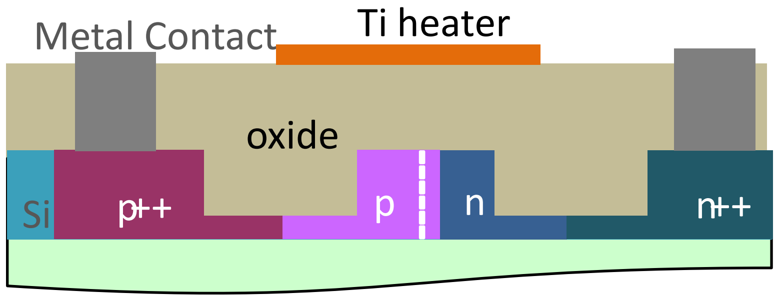


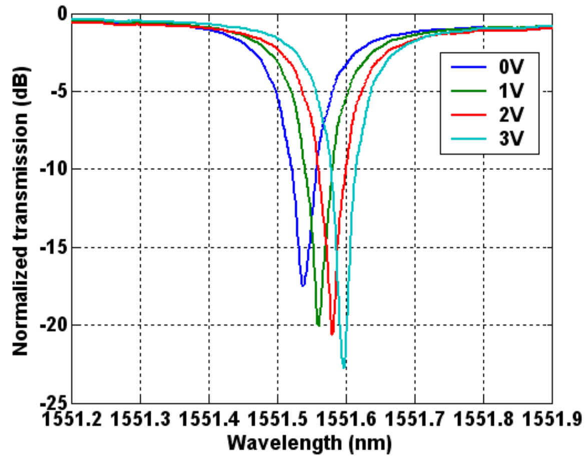


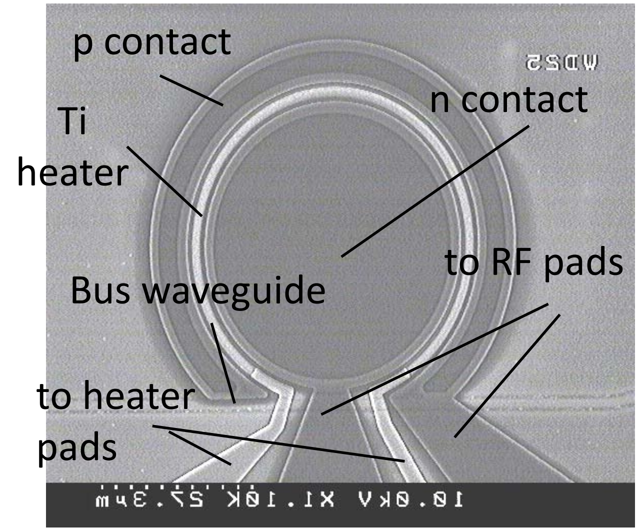
Krishnamoorthy et al., *Proceedings of the IEEE*, July 2009

25µm ring modulator w/ integrated heater

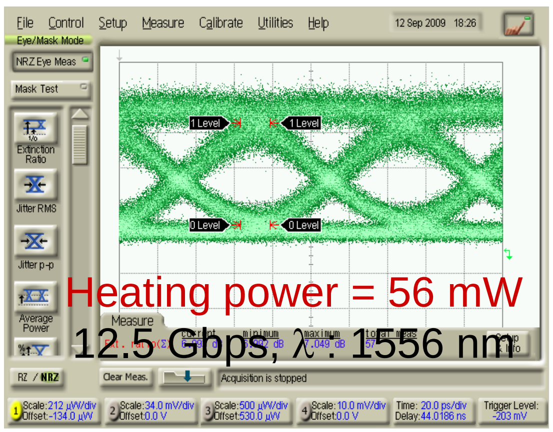
P. Dong et al., *IEEE Summer Topical Meeting on Optics in Data Centers*, July 2010

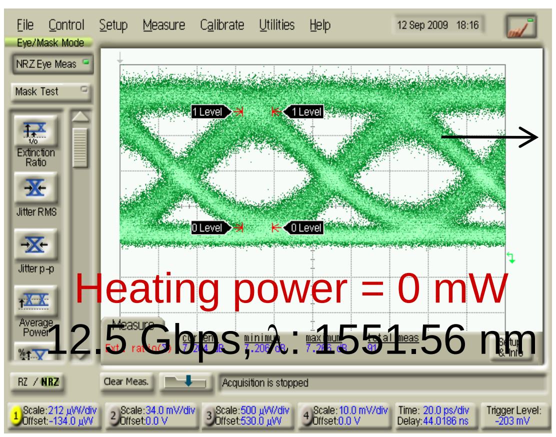


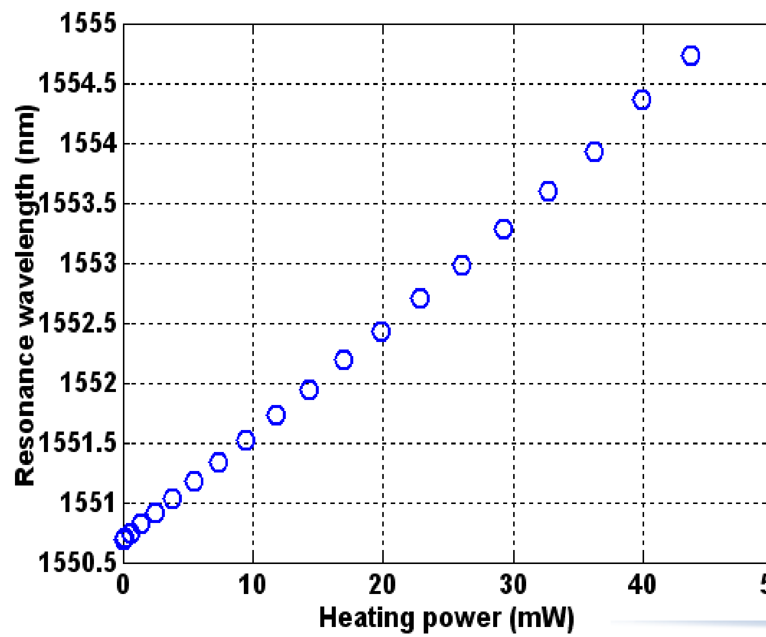




- Ring radius = 25 µm, FSR = 3.9 nm
- Total working wavelength range > 150 nm
- Heating power: **11.5 mW/nm, or 45 mW per FSR tuning**
- 12.5 Gbps is achieved with a V_{pp} = 3 V, RE>6dB, limited by BERT
- Modulation energy/power of **200 fJ/bit** or 2.5 mW
- No performance degradation over one FSR tuning

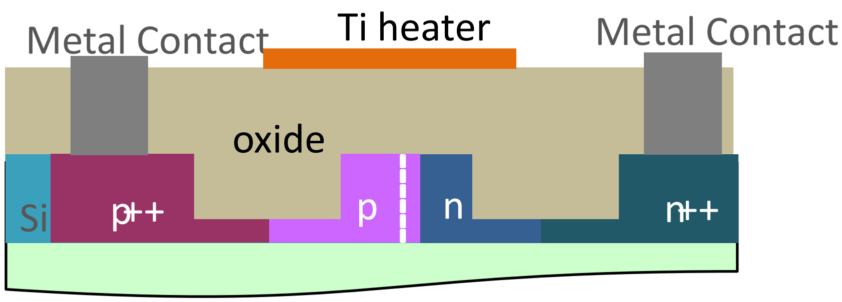


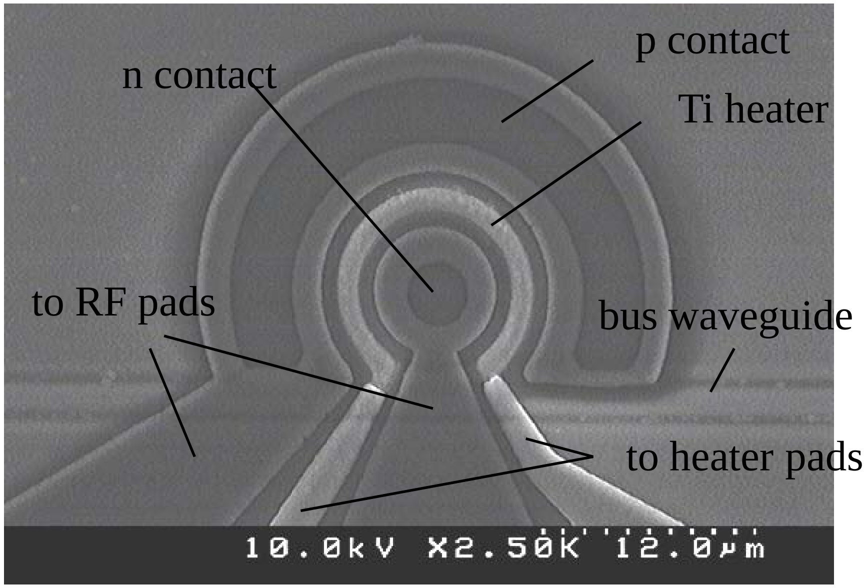




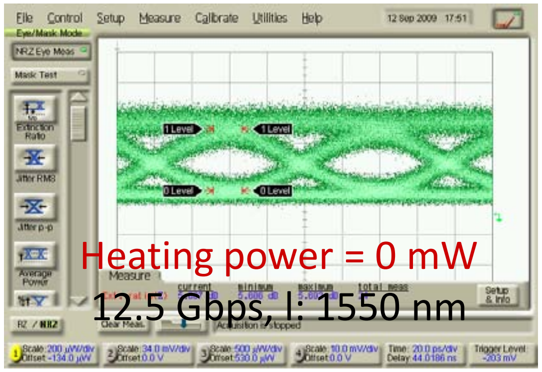
5µm ring modulator w/ integrated heater

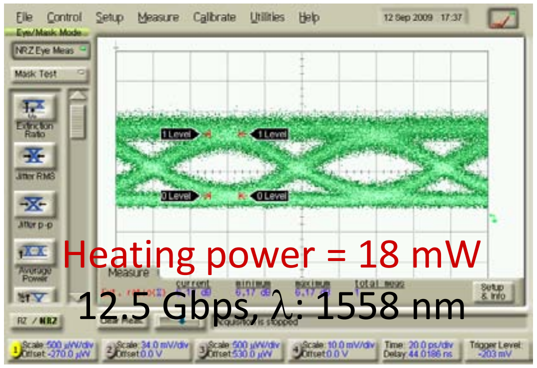
P. Dong et al., *Optics Express*, May 2010

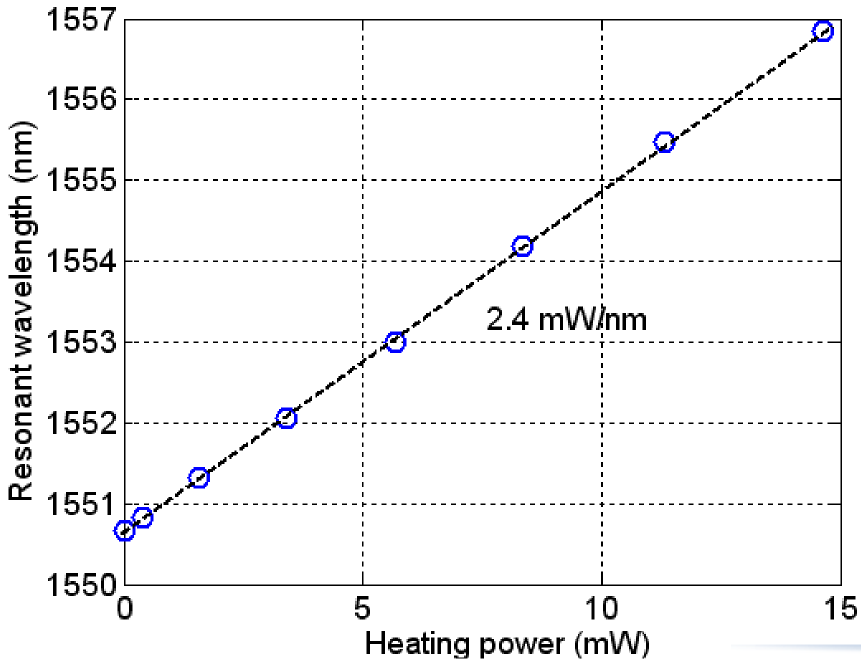




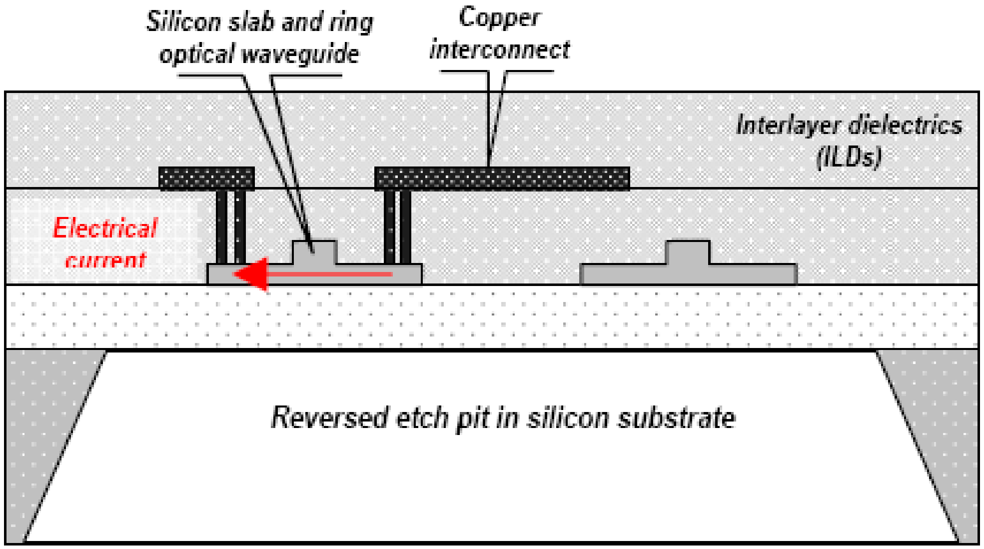
- **Ring radius = 5 µm**, FSR = 19 nm
- Heating efficiency**: 2.4 mW/nm, or 46 mW per FS**R tuning
- 12.5 Gbps is achieved with a V_{pp} = 3 V, 6dB ER
- Modulation energy of **40 fJ/bit** or 0.5 mW
- no detrimental effects found ver ~93 ºC range

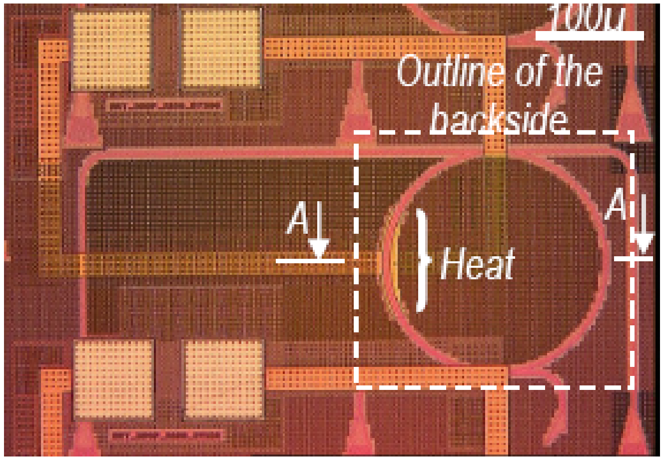


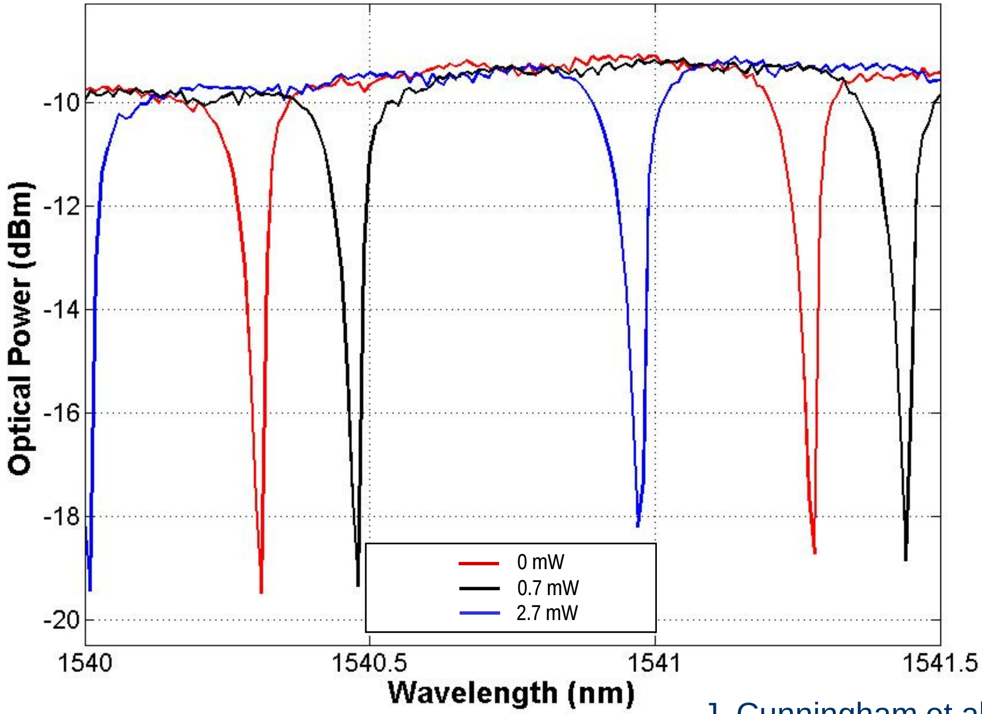




Efficient, tunable CMOS rings







- Add/drop tunable filters
- 0.13 µm SOI 6 metal CMOS process
- Dual heater stages with integrated waveguide heating
- Integrated back-side etch pit

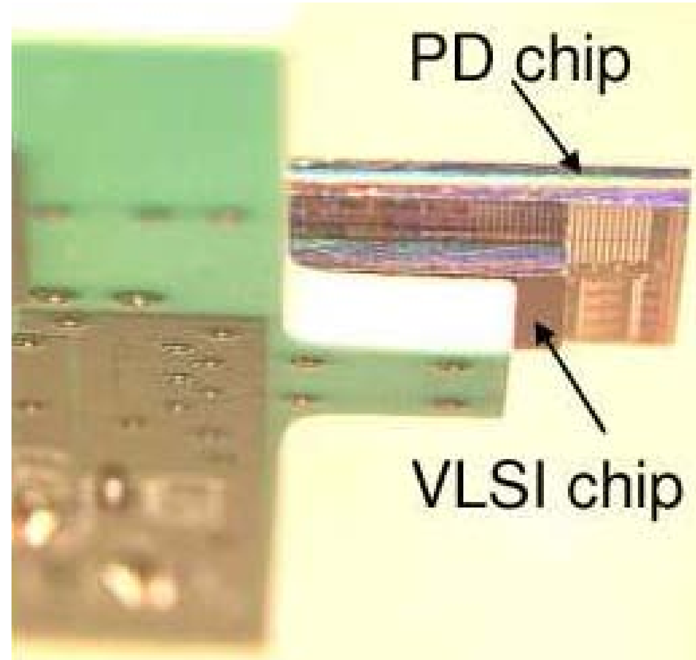
3.9 mW per 2π phase shift

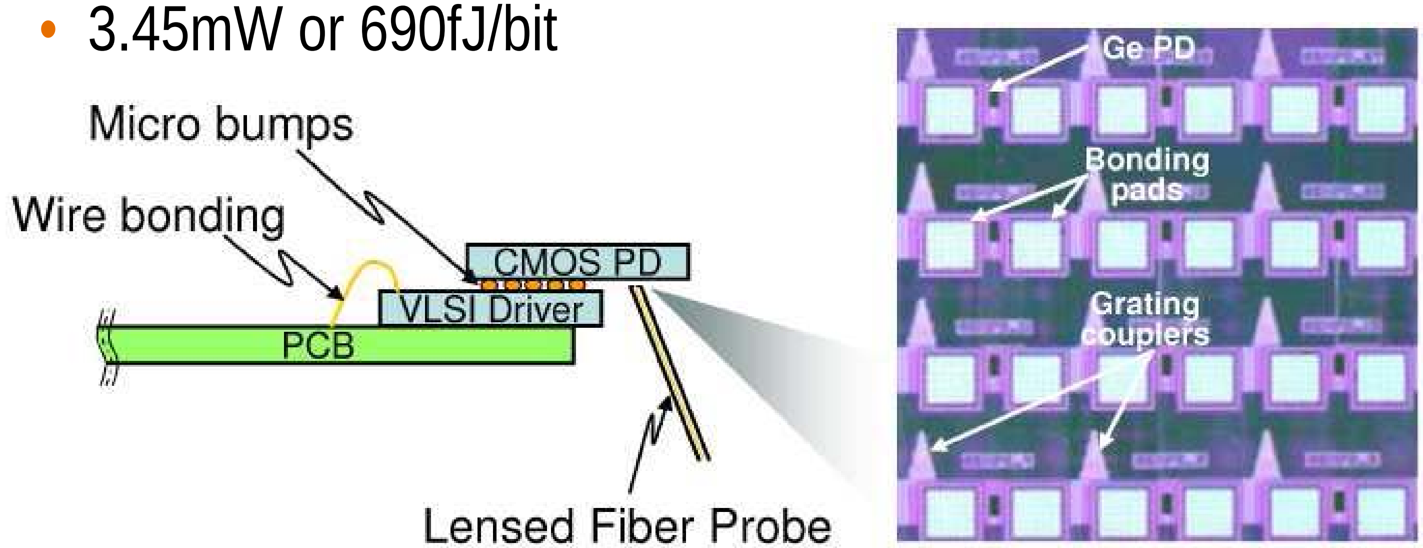
J. Cunningham et al., *IEEE Summer Top. Meet. on Optics in Data Centers*, July 2010

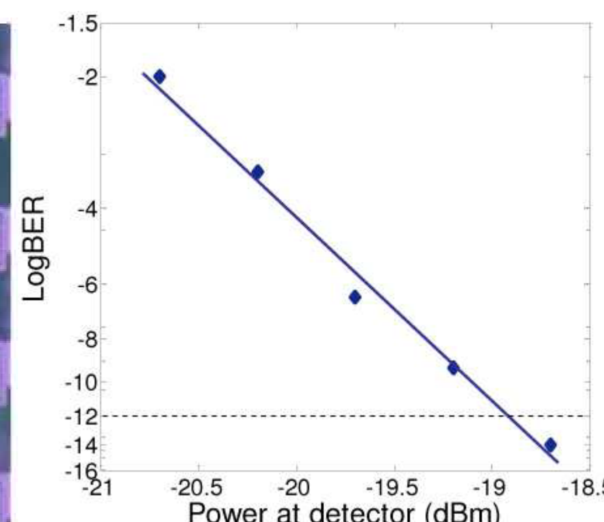
690fJ/bit all-CMOS Rx (circuits + device)

Performance Summary:

- CMOS integrated germanium photodetector
- 5Gbps, digitally clocked TIA-based receiver
- -18.9 dBm sensitivity at 10^{-12} BER
  with 0.7A/W responsivity, >10GHz BW, and
  <20fF detector capacitance
- BER measured below 10^{-14}
- 3.45mW or 690fJ/bit



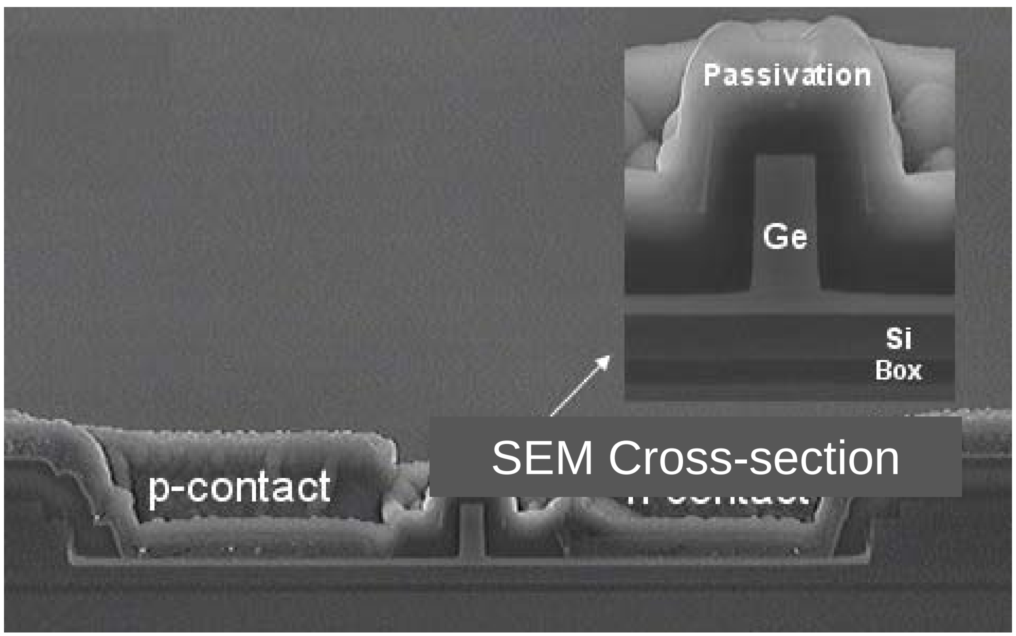




X. Zheng et al., *Optics Express*, January 2010

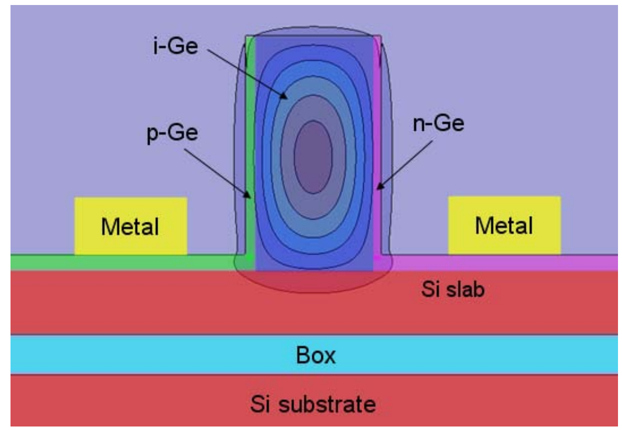
High-responsivity photodetectors

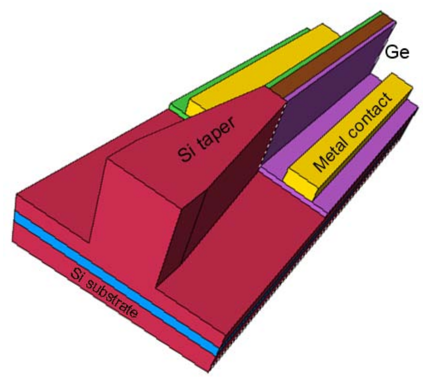
- Kotura Ge photodetectors
  - Butt-coupling between SOI and Ge waveguides enables short device lengths (~ 10 μm)
  - => Capacitance a few fF, device not RC-time limited)
  - Horizontal p-i-n junction design enables compatibility with larger Si waveguides
  - Narrow Ge WG width (0.65 μm) minimizes transit-time limitation (speed > 40 GHz)

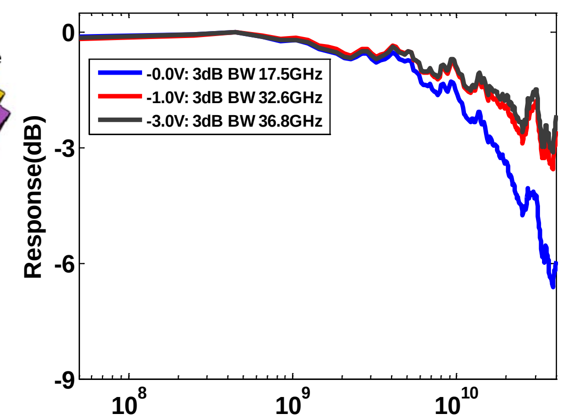


Performance Summary:

- Responsivity of 1.1 A/W @ 1550 nm
- Dark current of 0.24 μA @ -0.5 V, 1.3 μA @ -1 V
- Bandwidth > 32 GHz

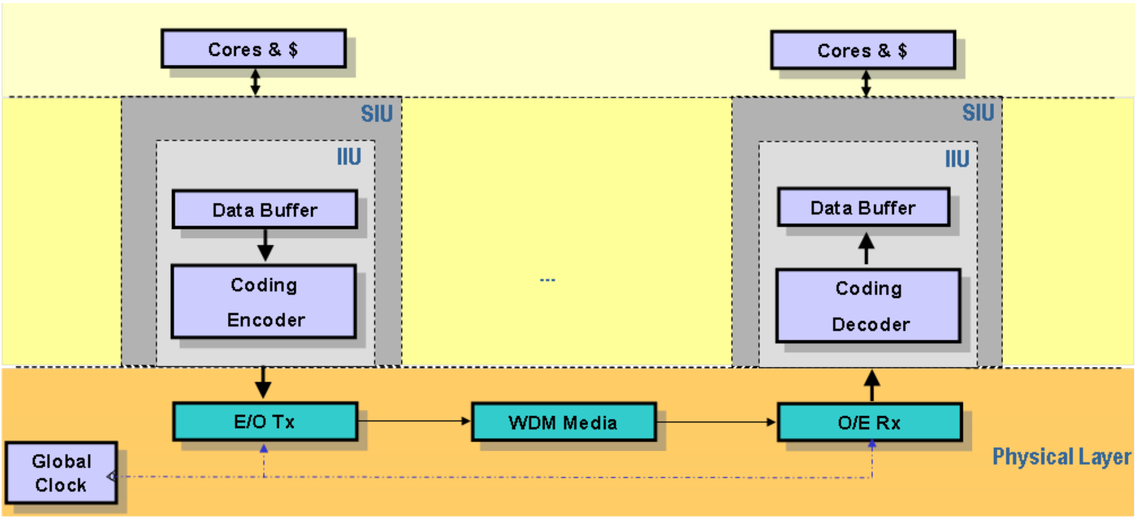


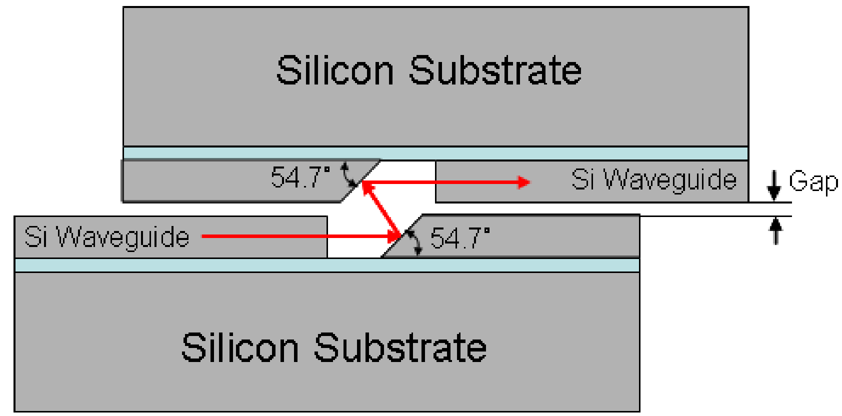




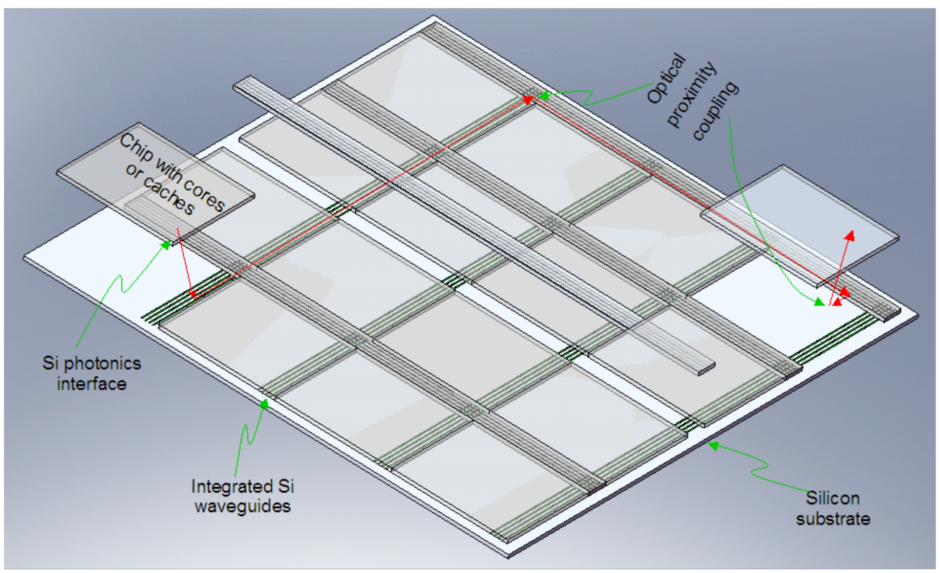
D. Feng et al., *Applied Physics Lett*, December 2009

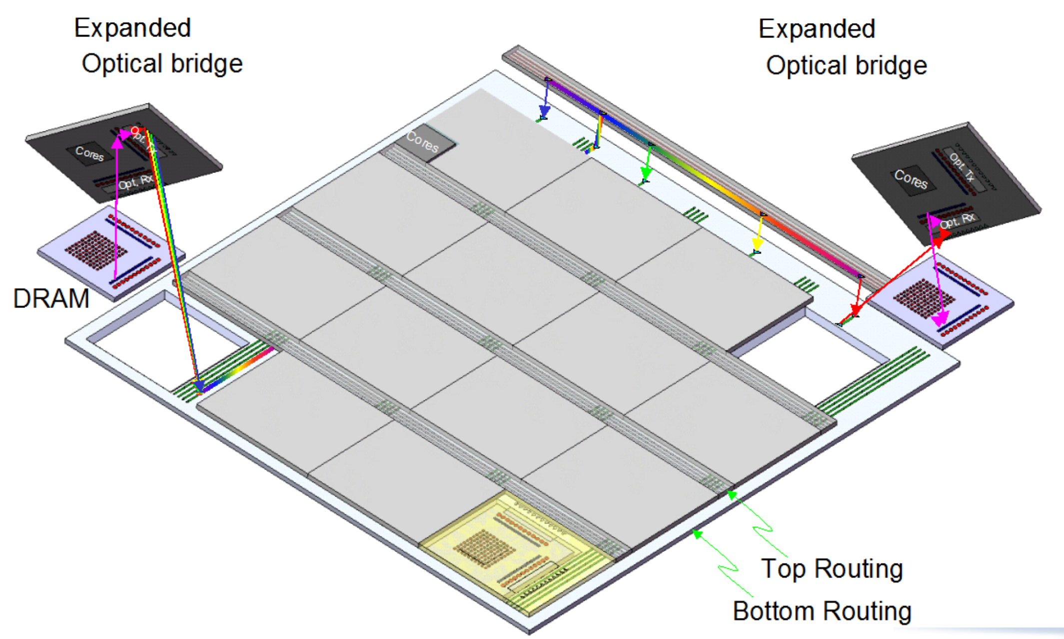
“Macrochip” logical & physical views



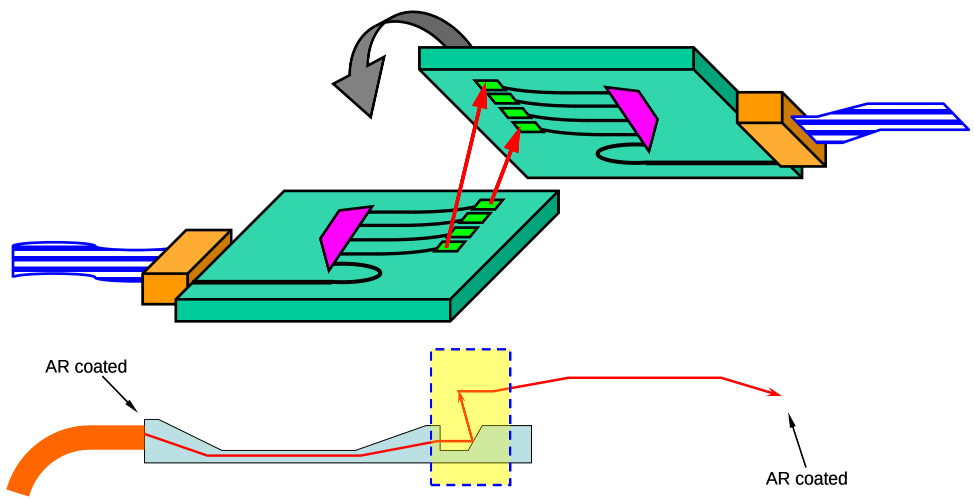


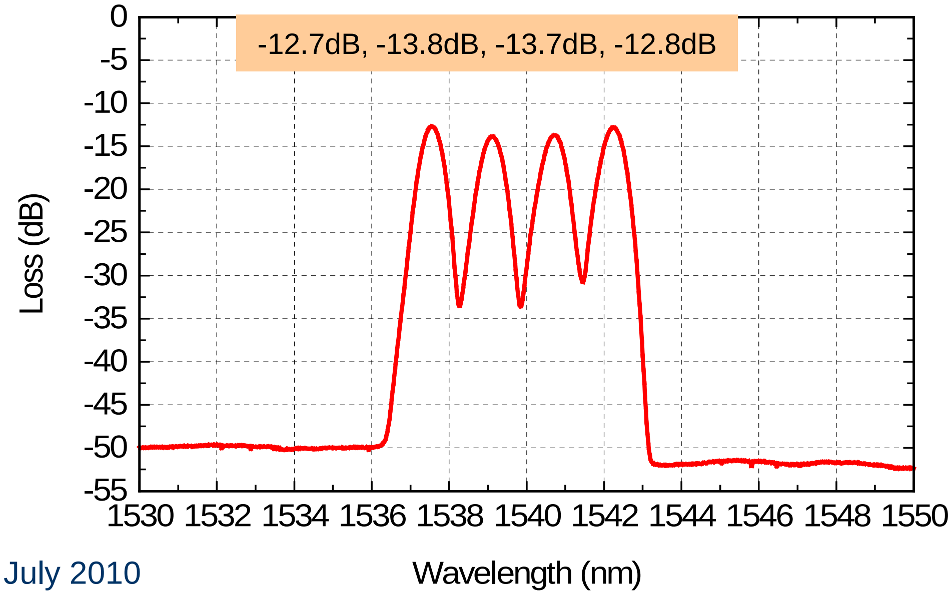
Krishnamoorthy et al., *Proceedings of the IEEE*, July 2009
R. Ho et al., *IEEE Communications Mag.,* July/August 2010



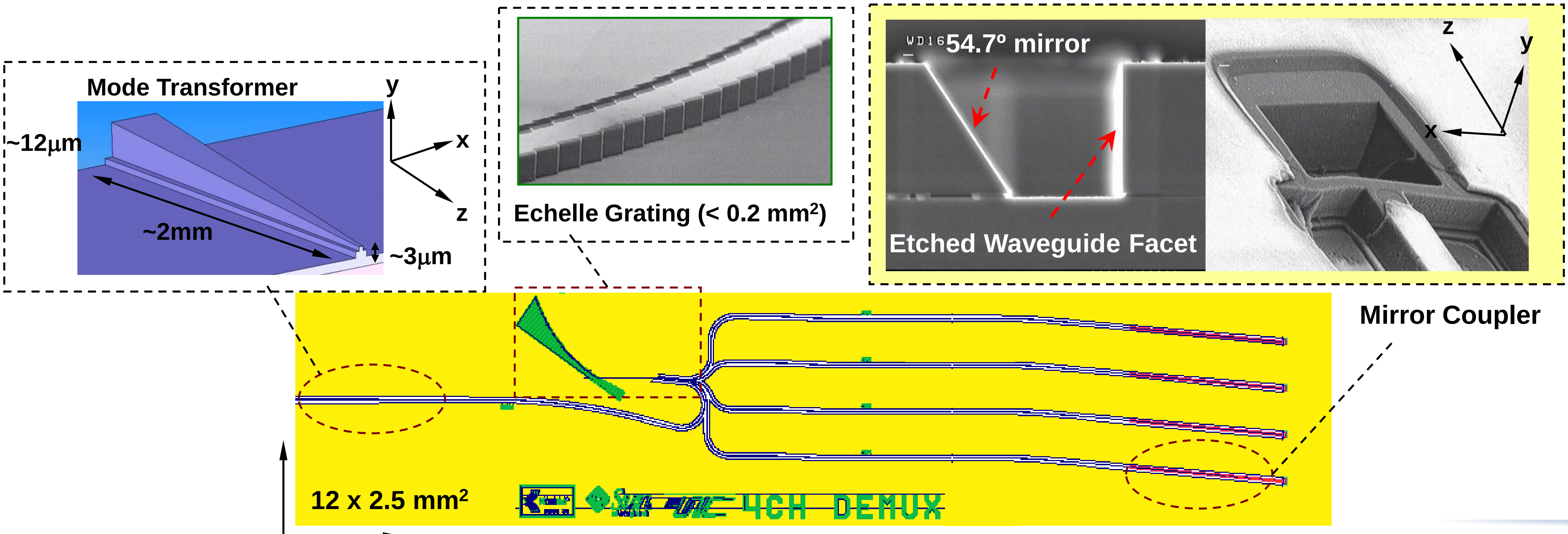


Interlayer link components



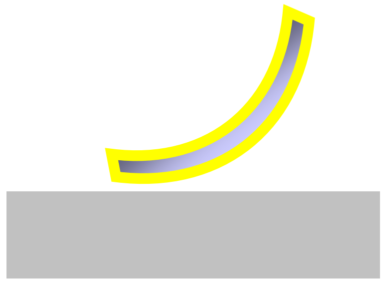


D. Lee et al., *IEEE Summer Topical Meeting on Optics in Data Centers*, July 2010

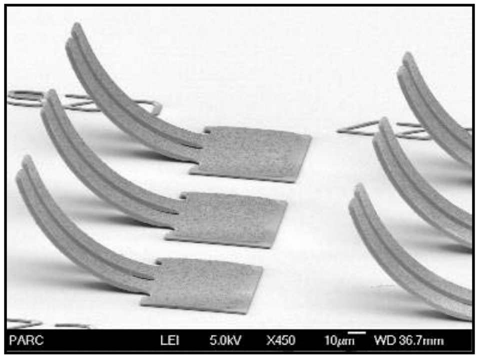


Rematable power, ground, & alignment

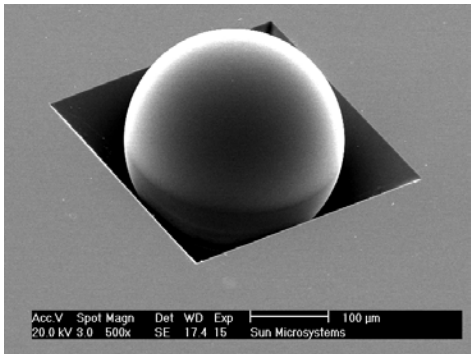
Sacrificial layer etch, spring lift-off and Au-plating

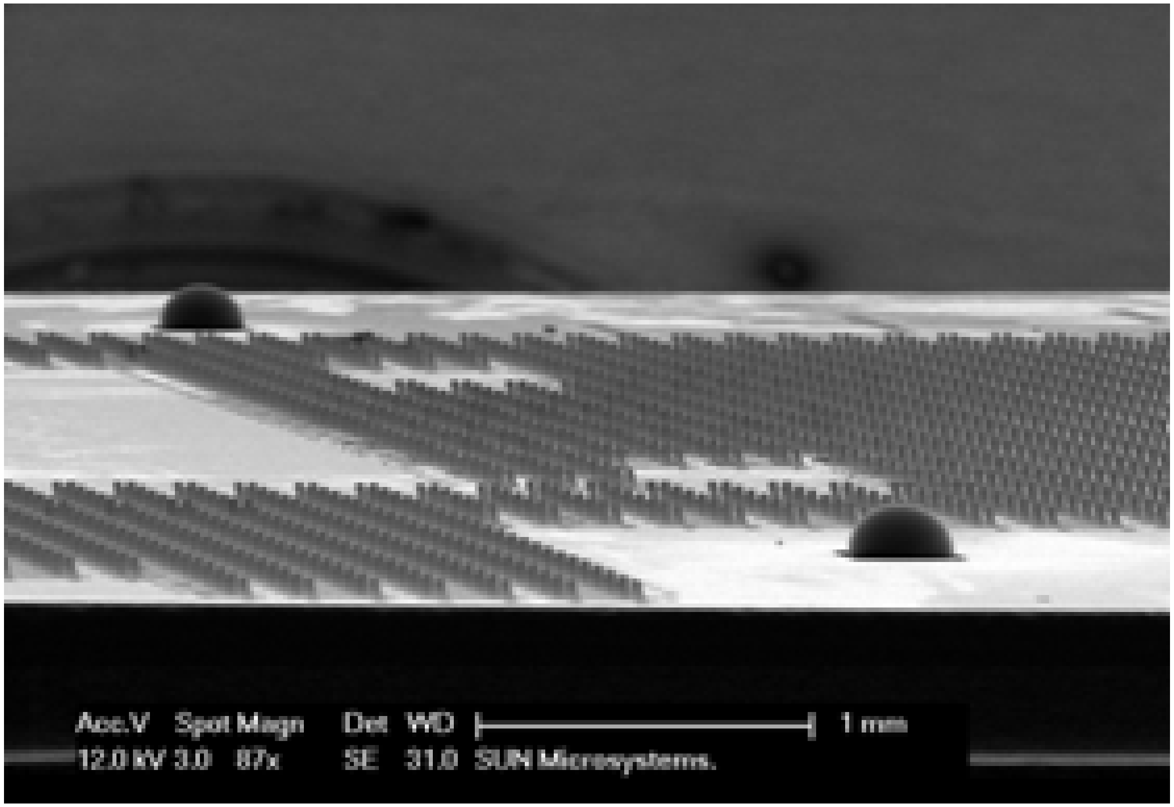


Micro-spring interconnects

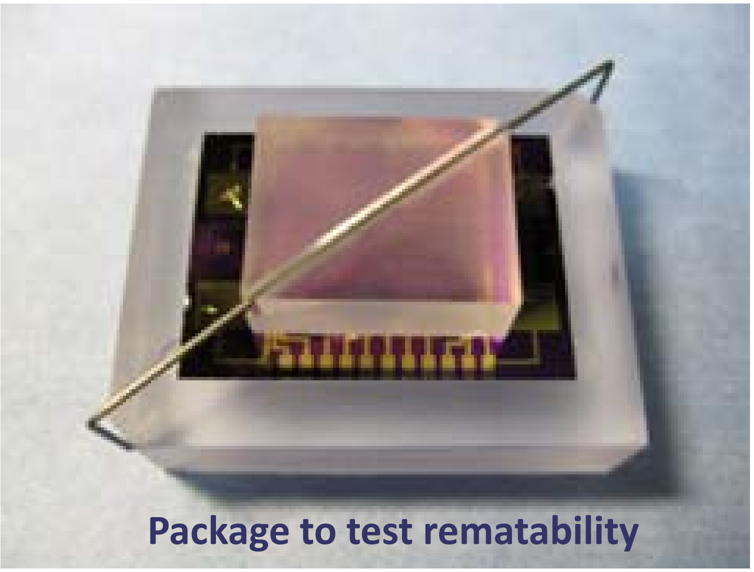


+





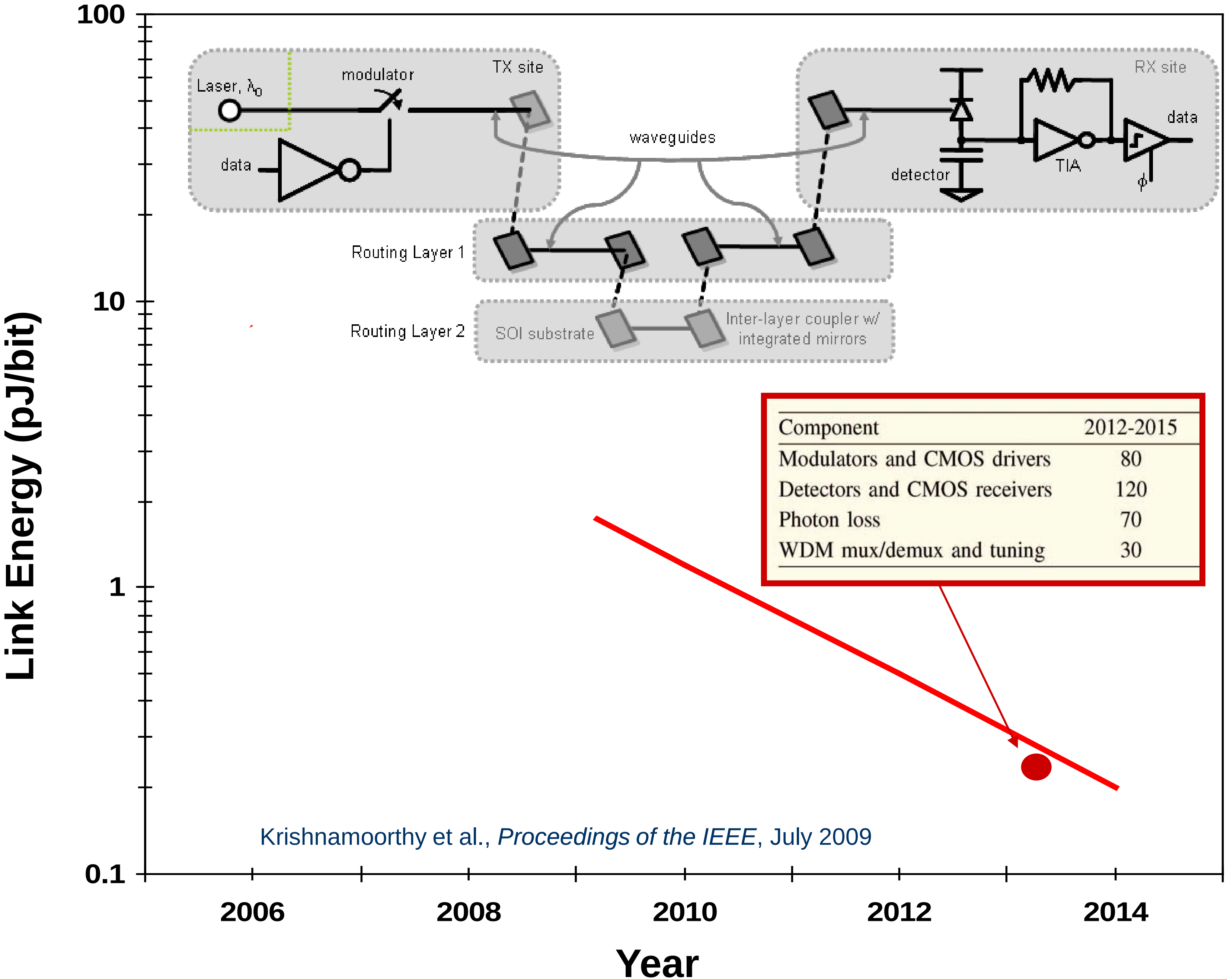
Co-integrate both technologies



Package to test rematability

I. Shubin et al., *IEEE ECTC*, May 2009

Optical interconnect energy roadmap



Component	2012-2015
Modulators and CMOS drivers	80
Detectors and CMOS receivers	120
Photon loss	70
WDM mux/demux and tuning	30

REFERENCES:

X. Zheng et al., *Optics Express,* October 2008

Krishnamoorthy et al., *IEEE JQE,* April 2009

I. Shubin et al., *Proc. ECTC,* May 2009

Krishnamoorthy et al., *Proc. of the IEEE,* July 2009

R. Ho et al., *IEEE ASSCC,* November 2009

D. Feng et al., *Applied Physics Lett.,* December 2009

X. Zheng et al., *Optics Express,* January 2010

X. Zheng et al., *Optics Express,* February 2010

J. Cunningham et al., *IEEE Photon. Summer Top. OND, TuD3.4,* July 2010

P. Dong et al., *IEEE Photonics Summer Top. OND, MD2.3,* July 2010

D. Lee et al., *IEEE Photonics Summer Top. OND, TuD3.3,* July 2010

R. Ho et al., *IEEE Communications Mag.,* July/August 2010

UNIC technology highlights to date

- Demonstration of passively-aligned multi-chip, multi-channel optical proximity communication
- Integration of ball-in-pit alignment with CMOS
- Record low-power silicon photonic link components
  - 320fJ/bit photonic Tx @ 5Gbps(w/ Kotura ring & custom driver)
  - 690fJ/bit photonic Rx @ 5Gbps (w/ Luxtera Ge PD & custom receiver)
  - 3.9mW FSR tunable mux/demux (w/ Luxtera ring & backside etch pit)
  - 1.1A responsivity, 0.24μA dark current large-core Ge detector (Kotura)
  - Areal density of ~730Gbps/sq. mm based on WDM link components
- Record efficiency SOI passive components
  - Thin silicon routing waveguide with 0.27dB/cm loss (Kotura)
  - 1x2 splitter with 0.1dB excess loss (Kotura)
- Demonstration of rematable power/gnd & chip alignment